```
;  PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1162
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-216-645-2

  Query Match            93.6%;  Score 5892.5;  DB 4;  Length 1162;
  Best Local Similarity  96.8%;  Pred. No. 0;
  Matches 1126;  Conservative    5;  Mismatches   17;  Indels   15;  Gaps    2;

Qy         52 MEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLAL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLAL 60

Qy        112 LLHEYYGASLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVP-- 169
              |||||||||||||||||||||||||||||||||| :||||||||||||||||||||
Db         61 LLHEYYGASLAWATSISIFENMNLRTLSEKARDDMK-NSPEDPEATMTDQGPSKEKVPEN 119

Qy        170 ------------GISQAVQQDSATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEG 217
                          || :  :      :   ||||||||||||||||||||||||||||||
Db        120 KYGMTKLILGVSDISDSNNKHKYVGIHSSFAEISQAMEQEGATAAETEEQEISQAMEQEG 179

Qy        218 ATAAETEEQGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGF 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATAAETEEQGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGF 239

Qy        278 RPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFIS 337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 RPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFIS 299

Qy        338 REWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRK 397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 REWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRK 359

Qy        398 VLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRA 457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 VLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRA 419

Qy        458 IINNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVR 517
              |:||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        420 IMNNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVR 479

Qy        518 RCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCE 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 RCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCE 539

Qy        578 EYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFL 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 EYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFL 599

Qy        638 FGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDK 697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 FGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDK 659

Qy        698 EFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVP 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 EFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVP 719

Qy        758 LWMRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMF 817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 LWMRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMF 779

Qy        818 RNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCG 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 RNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCG 839

Qy        878 LTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITA 937
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        840 LTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVMPLSDALRVSQCALQKLILEDCGITA 899

Qy        938 TGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGC 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 TGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGC 959

Qy        998 GSLALALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESL 1057
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GFLALALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESL 1019

Qy       1058 SCVISRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLA 1117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 SCVISRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLA 1079

Qy       1118 LSCNRHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLL 1177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 LSCNRHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLL 1139

Qy       1178 KPRVVIDGSWHSFDEDDRYWWKN 1200
              |||||||||||||||||||||||
Db       1140 KPRVVIDGSWHSFDEDDRYWWKN 1162


RESULT 5
US-10-416-642-1
; Sequence 1, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  RAMKUMAR, Jayalaxmi
;  APPLICANT:  ARVIZU, Chandra
;  TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
;  FILE REFERENCE: PF-0842 PCT
;  CURRENT APPLICATION NUMBER: US/10/416,642
;  CURRENT FILING DATE:  2003-05-13
;  PRIOR APPLICATION NUMBER: 60/249,407
;  PRIOR FILING DATE: 2000-11-15
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PERL Program
; SEQ ID NO 1
;   LENGTH: 1162
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CD1
US-10-416-642-1

  Query Match            93.6%;  Score 5892.5;  DB 4;  Length 1162;
  Best Local Similarity  96.8%;  Pred. No. 0;
  Matches 1126;  Conservative    5;  Mismatches   17;  Indels   15;  Gaps    2;

Qy         52 MEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLAL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLAL 60

Qy        112 LLHEYYGASLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVP-- 169
              |||||||||||||||||||||||||||||||||| :||||||||||||||||||||
Db         61 LLHEYYGASLAWATSISIFENMNLRTLSEKARDDMK-NSPEDPEATMTDQGPSKEKVPEN 119

Qy        170 ------------GISQAVQQDSATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEG 217
                          || :  :      :   ||||||||||||||||||||||||||||||
Db        120 KYGMTKLILGVSDISDSNNKHKYVGIHSSFAEISQAMEQEGATAAETEEQEISQAMEQEG 179

Qy        218 ATAAETEEQGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGF 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATAAETEEQGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGF 239

Qy        278 RPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFIS 337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 RPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFIS 299

Qy        338 REWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRK 397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        300 REWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRK 359

Qy        398 VLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRA 457

Db        360 VLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRA 419

Qy        458 IINNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVR 517

Db        420 IMNNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVR 479

Qy        518 RCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCE 577

Db        480 RCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCE 539

Qy        578 EYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFL 637

Db        540 EYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFL 599

Qy        638 FGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDK 697

Db        600 FGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDK 659

Qy        698 EFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVP 757

Db        660 EFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVP 719

Qy        758 LWMRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMF 817

Db        720 LWMRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMF 779

Qy        818 RNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCG 877

Db        780 RNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCG 839

Qy        878 LTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITA 937

Db        840 LTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVMPLSDALRVSQCALQKLILEDCGITA 899

Qy        938 TGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGC 997

Db        900 TGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGC 959

Qy        998 GSLALALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESL 1057

Db        960 GFLALALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESL 1019

Qy       1058 SCVISRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLA 1117

Db       1020 SCVISRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLA 1079

Qy       1118 LSCNRHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLL 1177

Db       1080 LSCNRHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLL 1139

Qy       1178 KPRVVIDGSWHSFDEDDRYWWKN 1200

Db       1140 KPRVVIDGSWHSFDEDDRYWWKN 1162
```

```
RESULT 6
US-10-216-645-4
; Sequence 4, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
;  APPLICANT: WEISS, BERTRAM
;  APPLICANT:  LESSL, MONIKA
;  APPLICANT:  PETERS-KOTTIG, MICHAELE
;  APPLICANT:  BECKMANN, GEORG
;  TITLE OF INVENTION: HUMAN MATER PROTEINS
;  FILE REFERENCE: SCH-1910
;  CURRENT APPLICATION NUMBER: US/10/216,645
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 1143
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-216-645-4

  Query Match            92.2%;  Score 5804;  DB 4;  Length 1143;
  Best Local Similarity  96.2%;  Pred. No. 0;
  Matches 1109;  Conservative   10;  Mismatches   20;  Indels   14;  Gaps    3;

Qy         52 MEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLAL 111

Db          1 MEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLAL 60

Qy        112 LLHEYYGASLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVP-- 169

Db         61 LLHEYYGASLAWATSISIFENMNLRTLSEKARDDMK-NSPEDPEATMTDQGPSKEKVPEN 119

Qy        170 --GISQAVQQDSATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQG 227

Db        120 KYGMTKLILGVS---------DISDSNNKHKYVGIHSSFAEISQAMEQEGATAAETEEQG 170

Qy        228 HGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGK 287

Db        171 HGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGK 230

Qy        288 SGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPV 347

Db        231 SGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPV 290

Qy        348 TEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIV 407

Db        291 TEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIV 350

Qy        408 TVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQ 467

Db        351 TVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRELLDQ 410

Qy        468 CQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVV 527

Db        411 CQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLEERVV 470

Qy        528 LKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSL 587

Db        471 LKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSL 530

Qy        588 QDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRR 647

Db        531 QDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRR 590

Qy        648 PLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSF 707

Db        591 PLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSF 650

Qy        708 QEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIE 767

Db        651 QEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIE 710

Qy        768 EQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQ 827

Db        711 EQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQ 770

Qy        828 HLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKIS 887

Db        771 HLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKIS 830

Qy        888 QILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASAL 947

Db        831 QILTTSPSLKSLSLAGNKVTDQGVMPLSDALRVSQCALQKLILEDCGITATGCQSLASAL 890
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:39:46 ; Search time 1.67987 Seconds
(without alignments)
465.507 Million cell updates/sec

Title: US-10-677-943A-4
Perfect score: 1379
Sequence: 1 LASALVSNRSLTHLCLSNNS.....VVIDGSWHSFDEDDRYWWKN 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	416	30.2	1062	7	US-11-137-465-43	Sequence 43, Appl
2	385.5	28.0	1049	7	US-11-137-465-42	Sequence 42, Appl
3	91.5	6.6	592	7	US-11-135-855-24	Sequence 24, Appl
4	88.5	6.4	437	1	US-10-632-150-54	Sequence 54, Appl
5	87.5	6.3	620	1	US-10-131-826A-324	Sequence 324, App
6	84.5	6.1	621	1	US-10-632-150-56	Sequence 56, Appl
7	82.5	6.0	1196	1	US-10-667-295-100	Sequence 100, App
8	81	5.9	483	7	US-11-137-465-40	Sequence 40, Appl
9	81	5.9	605	7	US-11-137-465-41	Sequence 41, Appl
10	78.5	5.7	642	1	US-10-131-826A-370	Sequence 370, App
11	76	5.5	674	1	US-10-131-826A-36	Sequence 36, Appl
12	75	5.4	4419	1	US-10-821-234-1155	Sequence 1155, Ap
13	74.5	5.4	483	1	US-10-632-150-48	Sequence 48, Appl
14	73	5.3	452	7	US-11-074-176-274	Sequence 274, App
15	71.5	5.2	421	1	US-10-858-730-1	Sequence 1, Appli
16	71	5.1	434	1	US-10-632-150-24	Sequence 24, Appl
17	71	5.1	1049	1	US-10-131-826A-358	Sequence 358, App
18	70	5.1	897	7	US-11-137-465-35	Sequence 35, Appl
19	70	5.1	993	7	US-11-137-465-36	Sequence 36, Appl
20	69.5	5.0	428	1	US-10-632-150-6	Sequence 6, Appli
21	69	5.0	283	1	US-10-632-150-50	Sequence 50, Appl
22	69	5.0	513	1	US-10-131-826A-536	Sequence 536, App
23	69	5.0	1170	1	US-10-831-997-2	Sequence 2, Appli
24	69	5.0	1170	7	US-11-046-456-28	Sequence 28, Appl
25	69	5.0	1170	7	US-11-046-644-28	Sequence 28, Appl
26	68.5	5.0	1032	1	US-10-835-475-1	Sequence 1, Appli
27	68.5	5.0	1032	7	US-11-014-367-1	Sequence 1, Appli
28	68	4.9	1036	1	US-10-131-826A-142	Sequence 142, App
29	67.5	4.9	555	7	US-11-034-569-8	Sequence 8, Appli
30	67.5	4.9	643	7	US-11-137-465-54	Sequence 54, Appl
31	67.5	4.9	5024	1	US-10-793-626-2964	Sequence 2964, Ap
32	67	4.9	945	1	US-10-131-826A-146	Sequence 146, App
33	66.5	4.8	585	1	US-10-821-234-1489	Sequence 1489, Ap
34	65	4.7	379	1	US-10-131-826A-328	Sequence 328, App
35	65	4.7	660	1	US-10-131-826A-350	Sequence 350, App
36	64.5	4.7	732	1	US-10-518-599-23	Sequence 23, Appl
37	64.5	4.7	1032	7	US-11-014-367-2	Sequence 2, Appli
38	64	4.6	373	7	US-11-082-389-248	Sequence 248, App
39	63.5	4.6	274	1	US-10-821-234-1284	Sequence 1284, Ap
40	63.5	4.6	1032	7	US-11-014-367-3	Sequence 3, Appli
41	63	4.6	152	7	US-11-034-569-4	Sequence 4, Appli
42	63	4.6	421	1	US-10-131-826A-302	Sequence 302, App
43	63	4.6	708	1	US-10-131-826A-298	Sequence 298, App
44	63	4.6	1338	1	US-10-821-234-1622	Sequence 1622, Ap
45	63	4.6	1338	7	US-11-109-156-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-11-137-465-43
; Sequence 43, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 1062
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-43

Query Match 30.2%; Score 416; DB 7; Length 1062;
Best Local Similarity 39.4%; Pred. No. 1.1e-33;
Matches 97; Conservative 38; Mismatches 111; Indels 0; Gaps 0;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              |: ||  |:||| : ||:| | :||  ||  ::| | | |||| |  |||  | |  ||
Db        803 LSLALEVNQSLTCVNLSDNELLDEGAKLLYTTLRHPKCFLQRLSLENCHLTEANCKDLAA 862

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
               |: :  |||| |: ||: : ||| ||| :| | | || | |  | :|:  |  |: ::
Db        863 VLVVSRELTHLCLAKNPIGNTGVKFLCEGLRYPECKLQTLVLWNCDITSDGCCDLTKLLQ 922

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
                 |  |||  | :|  |:  ||| |::    |  | |  | :    || |  |||||:
Db        923 EKSSLLCLDLGLNHIGVKGMKFLCEALRKPLCNLRCLWLWGCSIPPFSCEDLCSALSCNQ 982
```

```
Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              | :|:| ||     |:  |      | :  |: : |     :  :: |||||::   |:::
Db        983 SLVTLDDLGQNPLGSSGVKMLFETLTCSSGTLRTLRLKIDDFNDELNKLLEEIEEKNPQLI 1042

Qy        241 IDGSWH 246
              ||   |
Db       1043 IDTEKH 1048
```

RESULT 2
US-11-137-465-42
; Sequence 42, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-42

Query Match 28.0%; Score 385.5; DB 7; Length 1049;
Best Local Similarity 42.5%; Pred. No. 1.1e-30;
Matches 90; Conservative 27; Mismatches 94; Indels 1; Gaps 1;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||: || :: |||||: | :|: || || :  | | || |:|  |   |     |:|
Db        836 LAAVLVVSKKLTHLCLAKNPIGDTGVKFLCEGLSYPDCKLQTLVLVSCSATTQQWADLSL 895

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||  |  || ::|| | : | | |||   :| | | || | |  ||| | |: |: |:
Db        896 ALEVNQSLTCVNLSDNELLDEGAKLLYTTLRHPKCFLQRLSLENCHLTEANCKDLAAVLV 955

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              || |  | |  | :|: ||  |||||:     |  | |: | :|   |  || ||
Db        956 VSRELTHLCLAKNPIGNTGVKFLCEGLRYPECKLQTLVLQQCSITKLGCRYLSEALQEAC 1015

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQ 212
              ||:|:|  |  : :|:  || |   |  ||:
Db       1016 SLTNLDLSINQIA-RGLWILCQALENPNCNLK 1046
```

RESULT 3
US-11-135-855-24
; Sequence 24, Application US/11135855
; Publication No. US20050255557A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM p.l.c.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50013
; CURRENT APPLICATION NUMBER: US/11/135,855
; CURRENT FILING DATE: 2005-05-24
; PRIOR APPLICATION NUMBER: US/10/203,708
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/04703
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,172
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/186,084
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-135-855-24

Query Match 6.6%; Score 91.5; DB 7; Length 592;
Best Local Similarity 26.0%; Pred. No. 0.041;
Matches 60; Conservative 20; Mismatches 92; Indels 59; Gaps 9;

```
Qy         10 SLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLT 69
              :|  | || |:: :            ||   : ||  ||  |   |         : |  ||
Db         79 ALEELDLSENAIAHVEPGAFA---NLPRLRVLRLRGNQLKLIPPG----VFTRLDN--LT 129

Qy         70 HLSLSMN------------------PVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAA 110
              | || |                    | || :  :           |::| | :|:|||
Db        130 LLDLSENKLVILLDYTFQDLHSLRRLEVGDNDLVFVSRRAFAGLLALEELTLERCNLTAL 189

Qy        111 CCESLSCVIS----RSRHLKSLDLTDN----------------ALGDGGVAALCEGLKQK 150
               |||  : |    | |||    | |                  | :   |    ||
Db        190 SGESLGHLRSLGALRLRHLAIASLEDQNFRRLPGLLHLEIDNWPLLEEVAAGSLRGLN-- 247

Qy        151 NSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFS--PKGMMK 199
              || | :    :|:      : ||    ||| |||  |  |  |:|  :
Db        248 ---LTSLSVTHTNITA----VPAAALRHQAHLTCLNLSHNPISTVPRGSFR 291
```

RESULT 4
US-10-632-150-54
; Sequence 54, Application US/10632150
; Publication No. US20050251871A1
; GENERAL INFORMATION:
; APPLICANT: Chiaur, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/10/632,150
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/385,219
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: all Xaa positions
; OTHER INFORMATION: Xaa=unknown amino acid residue
US-10-632-150-54

Query Match 6.4%; Score 88.5; DB 1; Length 437;
Best Local Similarity 24.5%; Pred. No. 0.056;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:26:25 ; Search time 12.459 Seconds
(without alignments)
1712.037 Million cell updates/sec

Title: US-10-677-943A-4
Perfect score: 1379
Sequence: 1 LASALVSNRSLTHLCLSNNS.....VVIDGSWHSFDEDDRYWWKN 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	436	31.6	1034	2	US-10-183-770A-18	Sequence 18, Appl
2	429	31.1	603	2	US-10-781-294-64	Sequence 64, Appl
3	429	31.1	1035	2	US-10-781-294-24	Sequence 24, Appl
4	420.5	30.5	375	2	US-10-781-294-39	Sequence 39, Appl
5	420.5	30.5	994	2	US-10-781-294-16	Sequence 16, Appl
6	418.5	30.3	456	1	US-08-910-731-2	Sequence 2, Appli
7	418.5	30.3	456	1	US-08-795-395-2	Sequence 2, Appli
8	418.5	30.3	456	2	US-10-188-495-64	Sequence 64, Appl
9	410.5	29.8	456	1	US-08-910-731-8	Sequence 8, Appli
10	397	28.8	456	1	US-08-910-731-4	Sequence 4, Appli
11	397	28.8	456	1	US-08-795-395-4	Sequence 4, Appli
12	396.5	28.8	1033	2	US-10-029-347-4	Sequence 4, Appli
13	375.5	27.2	461	1	US-08-910-731-6	Sequence 6, Appli
14	374.5	27.2	516	2	US-10-104-047-2033	Sequence 2033, Ap
15	356	25.8	1429	2	US-10-029-347-3	Sequence 3, Appli
16	356	25.8	1429	2	US-10-183-770A-3	Sequence 3, Appli
17	354	25.7	1429	2	US-10-029-347-26	Sequence 26, Appl
18	354	25.7	1429	2	US-10-183-770A-15	Sequence 15, Appl
19	354	25.7	1429	2	US-09-996-617-2	Sequence 2, Appli
20	354	25.7	1473	2	US-09-388-221B-2	Sequence 2, Appli
21	350	25.4	1454	2	US-09-388-221B-10	Sequence 10, Appl
22	347	25.2	1399	2	US-09-388-221B-4	Sequence 4, Appli
23	347	25.2	1443	2	US-09-388-221B-6	Sequence 6, Appli
24	343	24.9	1424	2	US-09-388-221B-12	Sequence 12, Appl
25	253.5	18.4	514	2	US-10-029-347-25	Sequence 25, Appl
26	253.5	18.4	514	2	US-10-183-770A-14	Sequence 14, Appl
27	164	11.9	483	2	US-09-904-615-154	Sequence 154, App
28	164	11.9	483	2	US-10-054-988-154	Sequence 154, App
29	164	11.9	1040	2	US-10-002-974-59	Sequence 59, Appl
30	164	11.9	1040	2	US-10-002-974-85	Sequence 85, Appl
31	162	11.7	1013	2	US-10-014-269-3	Sequence 3, Appli
32	162	11.7	1013	2	US-10-002-974-3	Sequence 3, Appli
33	162	11.7	1040	2	US-10-014-269-2	Sequence 2, Appli
34	162	11.7	1040	2	US-10-014-269-4	Sequence 4, Appli
35	162	11.7	1040	2	US-10-002-974-2	Sequence 2, Appli
36	162	11.7	1040	2	US-10-002-974-4	Sequence 4, Appli
37	162	11.7	1040	2	US-10-002-974-57	Sequence 57, Appl
38	162	11.7	1040	2	US-10-002-974-61	Sequence 61, Appl
39	162	11.7	1040	2	US-10-002-974-65	Sequence 65, Appl
40	162	11.7	1040	2	US-10-002-974-67	Sequence 67, Appl
41	162	11.7	1040	2	US-10-002-974-69	Sequence 69, Appl
42	162	11.7	1040	2	US-10-002-974-87	Sequence 87, Appl
43	162	11.7	1040	2	US-10-002-974-89	Sequence 89, Appl
44	162	11.7	1043	2	US-09-949-016-9080	Sequence 9080, Ap
45	161	11.7	1040	2	US-10-002-974-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-10-183-770A-18
; Sequence 18, Application US/10183770A
; Patent No. 6949363
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRBM1
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770A
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-183-770A-18

Query Match 31.6%; Score 436; DB 2; Length 1034;
Best Local Similarity 44.0%; Pred. No. 6.2e-37;
Matches 95; Conservative 35; Mismatches 86; Indels 0; Gaps 0;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              | | | :::||| | ||:|||||: |: :|| ::: | |:::|| | :| |    |  ::|
Db        731 LFSVLSTSQSLTELDLSDNSLGDPGMRVLCETLQHPGCNIRRLWLGRCGLSHECCFDISL 790

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              |  |  |  | || | : | |::|||  ::   |:|: | || | ||:|||: |: |:|
Db        791 VLSSNQKLVELDLSDNALGDFGIRLLCVGLKHLLCNLKKLWLVSCCLTSACCQDLASVLS 850

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              |  |  | : :||||| ||| |||  |    | :|||  |||| || |||  || |:
Db        851 TSHSLTRLYVGENALGDSGVAILCEKAKNPQCNLQKLGLVNSGLTSVCCSALSSVLSTNQ 910

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGL 216
              :|| | |  |   ||:  ||     |   ||:: |
Db        911 NLTHLYLRGNTLGDKGIKLLCEGLLHPDCKLQVLEL 946
```

RESULT 2
US-10-781-294-64
; Sequence 64, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.

; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-781-294-64

Query Match 31.1%; Score 429; DB 2; Length 603;
Best Local Similarity 41.7%; Pred. No. 1.6e-36;
Matches 90; Conservative 36; Mismatches 90; Indels 0; Gaps 0;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||:|| :| :|  | |  |:||: || |||: :| |:| || | | :| : :: |  |:
Db        251 LAAALCTNPNLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSSACEDLSA 310

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||: |  || : || | |   |: |||| :| | | || ::| || | :  |: :: |:
Db        311 ALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLQMIQLRKCQLESGACQEMASVLG 370

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
               : ||  |||| ||| | |:  ||:||:     |  | || | ||:  |: |:  || |:
Db        371 TNPHLVELDLTGNALEDLGLRLLCQGLRHPVCRLRTLWLKICRLTAAACDELASTLSVNQ 430

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGL 216
                |  |:|  |     |:: ||     ||  || : |
Db        431 SLRELDLSLNELGDLGVLLLCEGLRHPTCKLQTLRL 466
```

RESULT 3
US-10-781-294-24
; Sequence 24, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-24

Query Match 31.1%; Score 429; DB 2; Length 1035;
Best Local Similarity 41.7%; Pred. No. 3.4e-36;
Matches 90; Conservative 36; Mismatches 90; Indels 0; Gaps 0;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||:|| :| :|  | |  |:||: || |||: :| |:| || | | :| : :: |  |:
Db        679 LAAALCTNPNLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSSACEDLSA 738

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||: |  || : || | |   |: |||| :| | | || ::| || | :  |: :: |:
Db        739 ALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLQMIQLRKCQLESGACQEMASVLG 798

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
               : ||  |||| ||| | |:  ||:||:     |  | || | ||:  |: |:  || |:
Db        799 TNPHLVELDLTGNALEDLGLRLLCQGLRHPVCRLRTLWLKICRLTAAACDELASTLSVNQ 858

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGL 216
                |  |:|  |     |:: ||     ||  || : |
Db        859 SLRELDLSLNELGDLGVLLLCEGLRHPTCKLQTLRL 894
```

RESULT 4
US-10-781-294-39
; Sequence 39, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-781-294-39

Query Match 30.5%; Score 420.5; DB 2; Length 375;
Best Local Similarity 39.0%; Pred. No. 6.9e-36;
Matches 94; Conservative 41; Mismatches 105; Indels 1; Gaps 1;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||  | :|: ||:| :| | | : || ||| ::  |   |  |||  |||     |  ::
Db        124 LAGLLTNNKKLTYLNVSCNQL-DTGVPLLCEALCSPDTVLVYLMLAFCHLSEQCCEYISE 182

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              |: |  : :| || | ::| |:| ||| :: | | |  | |||| :||| || |:  :
Db        183 MLLRNKSVRYLDLSANVLKDEGLKTLCEALKHPDCCLDSLCLVKCFITAAGCEDLASALI 242

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
               :::|| | :  | :|| ||  ||  |   :  |  |||: ||||| ||: |:  |:|::
Db        243 SNQNLKILQIGCNEIGDVGVQLLCRALTHTDCRLEILGLEECGLTSTCCKDLASVLTCSK 302

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
               | |||  |     |:: || |    |   ||::|| |  :  : : ||   :   | :
Db        303 TLQQLNLTLNTLDHTGVVVLCEALRHPECALQVLGLRKTDFDEETQALLTAEEERNPNLT 362

Qy        241 I 241
               |
Db        363 I 363
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:38:41 ; Search time 39.4769 Seconds
(without alignments)
2730.707 Million cell updates/sec

Title: US-10-677-943A-4
Perfect score: 1379
Sequence: 1 LASALVSNRSLTHLCLSNNS.....VVIDGSWHSFDEDDRYWWKN 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1379	100.0	1200	4	US-10-399-443-24	Sequence 24, Appl
2	1379	100.0	1200	4	US-10-677-943-24	Sequence 24, Appl
3	1379	100.0	1200	5	US-10-860-761-4	Sequence 4, Appli
4	1373	99.6	1033	4	US-10-092-900A-348	Sequence 348, App
5	1373	99.6	1143	4	US-10-216-645-4	Sequence 4, Appli
6	1373	99.6	1162	4	US-10-216-645-2	Sequence 2, Appli
7	1373	99.6	1162	4	US-10-416-642-1	Sequence 1, Appli
8	1335	96.8	1344	4	US-10-124-498-6	Sequence 6, Appli
9	1335	96.8	1344	4	US-10-066-521-6	Sequence 6, Appli
10	1301	94.3	258	4	US-10-399-443-4	Sequence 4, Appli
11	1301	94.3	258	4	US-10-677-943-4	Sequence 4, Appli
12	902	65.4	1111	4	US-10-216-645-5	Sequence 5, Appli
13	850	61.6	1111	4	US-10-399-443-6	Sequence 6, Appli
14	850	61.6	1111	4	US-10-677-943-6	Sequence 6, Appli
15	587.5	42.6	1093	5	US-10-794-342-13	Sequence 13, Appl
16	538.5	39.1	924	4	US-10-357-820-52	Sequence 52, Appl
17	477.5	34.6	896	4	US-10-124-498-22	Sequence 22, Appl
18	477.5	34.6	896	4	US-10-066-521-22	Sequence 22, Appl
19	436	31.6	1032	5	US-10-882-761-33	Sequence 33, Appl
20	436	31.6	1034	4	US-10-127-516-5	Sequence 5, Appli
21	436	31.6	1034	4	US-10-027-629-5	Sequence 5, Appli
22	436	31.6	1034	4	US-10-028-374-18	Sequence 18, Appl
23	436	31.6	1034	4	US-10-132-967-5	Sequence 5, Appli
24	436	31.6	1034	4	US-10-183-770-18	Sequence 18, Appl
25	436	31.6	1034	4	US-10-264-958B-22	Sequence 22, Appl
26	436	31.6	1034	6	US-11-107-572-18	Sequence 18, Appl
27	435	31.5	287	3	US-09-799-983-4	Sequence 4, Appli
28	435	31.5	1004	4	US-10-108-260A-3161	Sequence 3161, Ap
29	429	31.1	565	4	US-10-467-397-9	Sequence 9, Appli
30	429	31.1	603	3	US-09-965-621-64	Sequence 64, Appl
31	429	31.1	603	4	US-10-407-866-64	Sequence 64, Appl
32	429	31.1	603	4	US-10-781-294-64	Sequence 64, Appl
33	429	31.1	1027	4	US-10-407-866-68	Sequence 68, Appl
34	429	31.1	1035	3	US-09-965-621-24	Sequence 24, Appl
35	429	31.1	1035	4	US-10-407-866-24	Sequence 24, Appl
36	429	31.1	1035	4	US-10-781-294-24	Sequence 24, Appl
37	429	31.1	1061	4	US-10-124-498-18	Sequence 18, Appl
38	429	31.1	1061	4	US-10-066-521-18	Sequence 18, Appl
39	420.5	30.5	375	3	US-09-965-621-39	Sequence 39, Appl
40	420.5	30.5	375	4	US-10-407-866-39	Sequence 39, Appl
41	420.5	30.5	375	4	US-10-781-294-39	Sequence 39, Appl
42	420.5	30.5	919	4	US-10-094-749-2718	Sequence 2718, Ap
43	420.5	30.5	919	4	US-10-161-493-4	Sequence 4, Appli
44	420.5	30.5	919	4	US-10-398-037-7	Sequence 7, Appli
45	420.5	30.5	994	3	US-09-965-621-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-10-399-443-24
; Sequence 24, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-443-24

Query Match 100.0%; Score 1379; DB 4; Length 1200;
Best Local Similarity 100.0%; Pred. No. 1.2e-124;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 1062

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 1122

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 1182

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||||||||||||||||||
```

```
Db       1183 IDGSWHSFDEDDRYWWKN 1200


RESULT 2
US-10-677-943-24
; Sequence 24, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
;  APPLICANT: The Government of the United States of America as
;  APPLICANT:  represented by the Secretary of the Department of Health and
;  APPLICANT:  Human Services
;  APPLICANT:  Nelson, Lawrence
;  APPLICANT:  Tong, Zhi-Bin
;  TITLE OF INVENTION: Human Gene Critical to Fertility
;  FILE REFERENCE: 4239-64790
;  CURRENT APPLICATION NUMBER: US/10/677,943
;  CURRENT FILING DATE:  2003-10-01
;  PRIOR APPLICATION NUMBER: 60/241,510
;  PRIOR FILING DATE: 2000-10-18
;  PRIOR APPLICATION NUMBER: PCT/US02/09776
;  PRIOR FILING DATE: 2002-03-29
;  PRIOR APPLICATION NUMBER: PCT/US01/10981
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
;   LENGTH: 1200
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-677-943-24

  Query Match           100.0%;  Score 1379;  DB 4;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-124;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 1062

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 1122

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 1182

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||||||||||||||||||
Db       1183 IDGSWHSFDEDDRYWWKN 1200


RESULT 3
US-10-860-761-4
; Sequence 4, Application US/10860761
; Publication No. US20040248775A1
; GENERAL INFORMATION:
;  APPLICANT: WYETH
;  TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS
;  FILE REFERENCE: AM101318
;  CURRENT APPLICATION NUMBER: US/10/860,761
;  CURRENT FILING DATE:  2004-06-03
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 1200
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-860-761-4

  Query Match           100.0%;  Score 1379;  DB 5;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-124;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 1062

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 1122

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 1182

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||||||||||||||||||
Db       1183 IDGSWHSFDEDDRYWWKN 1200


RESULT 4
US-10-092-900A-348
; Sequence 348, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Taupier Jr., Raymond J.
;  APPLICANT:  Pena, Carol E.A.
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Gusev, Vladimir Y.
;  APPLICANT:  Ji, Weizhen
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Tchernev,, Velizar T.
;  APPLICANT:  Fernandes, Elma R.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Liu, Yi
;  APPLICANT:  Anderson, David W.
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Catterton, Elina
;  APPLICANT:  Leite, Mario W.
;  APPLICANT:  Zhong, Haihong
;  APPLICANT:  Alsobrook, John P.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Catherine E.
;  TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-290C
;  CURRENT APPLICATION NUMBER: US/10/092,900A
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: USSN 60/274,322
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/283,675
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: USSN 60/338,092
;  PRIOR FILING DATE: 2001-12-03
;  PRIOR APPLICATION NUMBER: USSN 60/274,281
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:22:26 ; Search time 9.93923 Seconds
(without alignments)
2497.570 Million cell updates/sec

Title: US-10-677-943A-4
Perfect score: 1379
Sequence: 1 LASALVSNRSLTHLCLSNNS.....VVIDGSWHSFDEDDRYWWKN 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	902	65.4	1111	2	A59000	mater protein [imp
2	418.5	30.3	456	2	A31857	ribonuclease inhib
3	397	28.8	456	2	S20597	ribonuclease inhib
4	375.5	27.2	461	2	A31858	ribonuclease-angio
5	354	25.7	1192	2	T17255	hypothetical prote
6	158.5	11.5	631	2	C89243	protein F28C1.3 [i
7	158.5	11.5	631	2	T21471	hypothetical prote
8	157.5	11.4	506	2	A45841	T-complex-associat
9	151	10.9	533	2	T52063	ran GTPase-activat
10	150	10.9	271	2	D84586	hypothetical prote
11	147	10.7	535	2	T48102	RAN GTPase activat
12	146.5	10.6	545	2	T52068	RAN GTPase-activat
13	141.5	10.3	483	2	S27880	Nasopressin recept
14	124.5	9.0	568	2	F86291	hypothetical prote
15	124	9.0	587	2	JC5300	Ran GTPase activat
16	124	9.0	589	2	A36983	RNA1 homolog fug1
17	122	8.8	589	2	T52070	RNA1 protein homol
18	122	8.8	1121	2	T02764	myosin-I binding p
19	116.5	8.4	183	2	F97803	hypothetical prote
20	115.5	8.4	789	2	T52067	hypothetical prote
21	114.5	8.3	489	2	S44609	hypothetical prote
22	113	8.2	360	2	A96799	hypothetical prote
23	110.5	8.0	1130	2	A48843	MHC class II trans
24	107.5	7.8	332	2	T52069	GTPase-activating
25	107.5	7.8	1411	2	S44770	C29E4.3 protein -
26	103.5	7.5	418	2	T16410	hypothetical prote
27	103	7.5	934	2	T05201	hypothetical prote
28	102	7.4	1088	2	F11A6.9 protein -
29	101.5	7.4	360	2	B86347	hypothetical prote
30	101.5	7.4	1232	2	T05322	hypothetical prote
31	101	7.3	607	2	E96598	protein F20N2.2 [i
32	99	7.2	1139	2	T08421	resistance protein
33	98.5	7.1	998	2	T23427	hypothetical prote
34	97	7.0	364	2	T08903	hypothetical prote
35	97	7.0	618	2	T48193	hypothetical prote
36	97	7.0	661	2	I56258	RP105 - mouse
37	94.5	6.9	538	2	T47624	hypothetical prote
38	94	6.8	367	2	AD0063	hypothetical prote
39	94	6.8	2233	2	T28669	surface protein 51
40	93.5	6.8	375	1	TDHUM4	monocyte surface g
41	92.5	6.7	612	2	T10727	protein kinase Xa2
42	92.5	6.7	628	2	A84649	probable glucose r
43	92.5	6.7	659	2	T01520	hypothetical prote
44	92.5	6.7	668	2	T05257	probable disease r
45	92	6.7	630	2	S13724	ESAG 8 protein - T

ALIGNMENTS

RESULT 1
A59000
mater protein [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: A59000
R;Tong, Z.B.; Nelson, L.M.
Endocrinology 140, 3720-3726, 1999
A;Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature c
A;Reference number: A59000; MUID:99360614; PMID:10433232
A;Accession: A59000
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1111 <KUR>
A;Cross-references: UNIPROT:Q9R1M5; UNIPARC:UPI0000021410; GB:AF074018; NID:g5802697; P
C;Genetics:
A;Gene: Mater

Query Match 65.4%; Score 902; DB 2; Length 1111;
Best Local Similarity 64.7%; Pred. No. 7.2e-70;
Matches 167; Conservative 37; Mismatches 54; Indels 0; Gaps 0;

```
Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              | ||| ||::||||||||||||| |||  ||: :| | |:||||:|| |::     | ||:
Db        854 LVSALFSNQNLTHLCLSNNSLGTEGVQQLCQFLRNPECALQRLILNHCNIVDDAYGFLAM 913

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
               |  |: ||||||:|||| |  :||||| ::||:|:||:|||| | ||  ||| |:|:|:
Db        914 RLANNTKLTHLSLTMNPVGDGAMKLLCEALKEPTCYLQELELVDCQLTQNCCEDLACMIT 973

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ::||||||| :||||| ||  ||||||| :| | |||| || |||:||||||||:|||
Db        974 TTKHLKSLDLGNNALGDKGVITLCEGLKQSSSSLRRLGLGACKLTSNCCEALSLAISCNP 1033

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              || |||||:|:||  ||:|||||| || ||| |||||| :|  ::|: ||||: :|| ||
Db       1034 HLNSLNLVKNDFSTSGMLKLCSAFQCPVSNLGIIGLWKQEYYARVRRQLEEVEPFVKPHVV 1093

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||| |:: ||||| ||||
Db       1094 IDGDWYASDEDDRNWWKN 1111
```

RESULT 2
A31857
ribonuclease inhibitor, hepatic - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: A31857; A35830
R;Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.
Biochemistry 27, 8537-8544, 1988
A;Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals t

```
A;Reference number: A31857; MUID:89118268; PMID:3219361
A;Accession: A31857
A;Molecule type: protein
A;Residues: 1-456 <HOF>
A;Cross-references: UNIPROT:P10775; UNIPARC:UPI000013392C
R;Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
Biochemistry 29, 8827-8834, 1990
A;Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclea
A;Reference number: A35830; MUID:91104783; PMID:2271559
A;Accession: A35830
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 82-456 <VIC>
A;Cross-references: UNIPARC:UPI000016C6F0; GB:M58700; NID:g164638; PIDN:AAA63448.1; PID:
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: liver
F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

  Query Match             30.3%;  Score 418.5;  DB 2;  Length 456;
  Best Local Similarity   34.7%;  Pred. No. 1.6e-28;
  Matches   95;  Conservative   43;  Mismatches  105;  Indels   31;  Gaps    3;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              : ||| :| ||| ||| | ||: ||:|: : :: | | :|:| |  | | ||||
Db         43 IGSALRANPSLTELCLRTNELGDAGVHLVLQGLQSPTCKIQKLSLQNCSLTEAGCGVLPS 102

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
                |    | | || ||: | |::|||| : :| |||: |:|  | |||| || |: |:
Db        103 TLRSLPTLRELHLSDNPLGDAGLRLLCEGLLDPQCHLEKLQLEYCRLTAASCEPLASVLR 162

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              :| || | :::| :|: |   | :||      |  | |: ||||   |: |   ::
Db        163 ATRALKELTVSNNDIGEAGARVLGQGLADSACQLETLRLENCGLTPANCKDLCGIVASQA 222

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPV----------QIRKL 228
              |  |:|  |     |: :||     | | |:   || |:  :           | ::
Db        223 SLRELDLGSNGLGDAGIAELCPGLLSPASRLKT--LWLWECDITASGCRDLCRVLQAKET 280

Qy        229 LEEV----------------QLLKPRVVIDGSW 245
              |:|:                ||:|   ::  |
Db        281 LKELSLAGNKLGDEGARLLCESLLQPGCQLESLW 314


RESULT 3
S20597
ribonuclease inhibitor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S20597
R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992
A;Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distributio
A;Reference number: S20597; MUID:92162755; PMID:1536887
A;Accession: S20597
A;Molecule type: mRNA
A;Residues: 1-456 <KAW>
A;Cross-references: UNIPROT:P29315; UNIPARC:UPI000013392D; EMBL:X62528; NID:g57670; PIDN
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

  Query Match             28.8%;  Score 397;  DB 2;  Length 456;
  Best Local Similarity   37.2%;  Pred. No. 1.2e-26;
  Matches   90;  Conservative   38;  Mismatches  106;  Indels    8;  Gaps    2;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              : ||: :| :|| | |  | ||: || |: : :: | | :|:| |  | | ||||
Db         43 IRSAIQANPALTELSLRTNELGDAGVGLVLQGLQNPTCKIQKLSLQNCSLTEAGCGVLPD 102

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
                |   | |  | |: ||: | |:|||| :|:| | |: |:|  |:|||  || |: |:
Db        103 VLRSLSTLRELHLNDNPLGDEGLKLLCEGLRDPQCRLEKLQLEYCNLTATSCEPLASVLR 162

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
                |    | | || ||: | |::|||| : :| |||: |:|  | |||| || |: |:
Db        163 VKPDFKELVLSNNDFHEAGIHTLCQGLKDSACQLESLKLENCGITSANCKDLCDVVASKA 222

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPV------QIRKLLEEVQL 234
              |  |:|  |     |: |||     |:  |:   || |   |         : ::|   |
Db        223 SLQELDLGSNKLGNTGIAALCSGLLLPSCRLRT--LWLWDCDVTAEGCKDLCRVLRAKQS 280

Qy        235 LK 236
              ||
Db        281 LK 282


RESULT 4
A31858
ribonuclease-angiogenin inhibitor - human
N;Alternate names: ribonuclease inhibitor, placental
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: A31858; S02012; S23933; S48636; T47188
R;Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
Biochemistry 27, 8545-8553, 1988
A;Title: Primary structure of human placental ribonuclease inhibitor.
A;Reference number: A31858; MUID:89118269; PMID:3219362
A;Accession: A31858
A;Molecule type: mRNA
A;Residues: 1-461 <LEE>
A;Cross-references: UNIPROT:P13489; UNIPARC:UPI0000000DE4; GB:M22414; NID:g186260; PIDI
R;Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.
EMBO J. 7, 4151-4156, 1988
A;Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) disclo
A;Reference number: S02012; MUID:89210799; PMID:3243277
A;Accession: S02012
A;Molecule type: mRNA
A;Residues: 1-422,'SE',425-461 <SCH>
A;Cross-references: UNIPARC:UPI000016AF5C; EMBL:X13973; NID:g35843; PIDN:CAA32151.1; P
A;Note: part of this sequence, including the carboxyl end of the mature protein, was c
R;Crevel-Thieffry, I.; Cotterill, S.; Schuller, E.
Biochim. Biophys. Acta 1122, 107-112, 1992
A;Title: Characterisation of a tryptic peptide from human placental ribonuclease inhib
A;Reference number: S23933; MUID:92338217; PMID:1633192
A;Accession: S23933
A;Molecule type: protein
A;Residues: 174-195 <CRE>
A;Cross-references: UNIPARC:UPI0000176402
R;Nadano, D.; Yasuda, T.; Takeshita, H.; Uchide, K.; Kishi, K.
Arch. Biochem. Biophys. 312, 421-428, 1994
A;Title: Purification and characterization of human brain ribonuclease inhibitor.
A;Reference number: S48636; MUID:94311593; PMID:8037455
A;Accession: S48636
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-14 <NAD>
A;Cross-references: UNIPARC:UPI0000176403
R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24374
A;Accession: T47188
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <AAA>
A;Cross-references: UNIPARC:UPI0000000DE4; EMBL:AL161967
A;Experimental source: adult testis; clone DKFZp434K249
C;Genetics:
A;Gene: GDB:RNH
A;Cross-references: GDB:125274; OMIM:173320
A;Map position: 11p15.5-11p15.5
A;Note: DKFZp434K249.1
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homolo
C;Keywords: blocked amino end; duplication
F;34-433/Region: leucine-rich 57-residue repeats
F;57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;174-195/Region: inhibitory
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:22:06 ; Search time 50.5361 Seconds
(without alignments)
3601.908 Million cell updates/sec

Title: US-10-677-943A-4
Perfect score: 1379
Sequence: 1 LASALVSNRSLTHLCLSNNS.....VVIDGSWHSFDEDDRYWWKN 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1379	100.0	1200	1	NALP5_HUMAN	P59047 homo sapien
2	907	65.8	1111	2	Q6VSG0_MOUSE	Q6vsg0 mus musculu
3	902	65.4	1095	2	Q6VSG5_MOUSE	Q6vsg5 mus musculu
4	902	65.4	1111	1	NALP5_MOUSE	Q9r1m5 mus musculu
5	902	65.4	1111	2	Q53ZD5_MOUSE	Q53zd5 mus musculu
6	902	65.4	1111	2	Q7TPU9_MOUSE	Q7tpu9 mus musculu
7	899	65.2	1095	2	Q6VSG3_MOUSE	Q6vsg3 mus musculu
8	899	65.2	1111	2	Q6XZF2_MOUSE	Q6xzf2 mus musculu
9	879.5	63.8	1098	2	Q647I9_BOVIN	Q647i9 bos taurus
10	771	55.9	1085	2	Q6VSG1_MOUSE	Q6vsg1 mus musculu
11	717	52.0	269	2	Q9D458_MOUSE	Q9d458 mus musculu
12	587.5	42.6	1093	1	NAL14_HUMAN	Q86w24 homo sapien
13	484.5	35.1	986	2	Q66X14_MOUSE	Q66x14 mus musculu
14	484.5	35.1	993	2	Q6B966_MOUSE	Q6b966 mus musculu
15	449	32.6	996	2	Q4U4S5_BOVIN	Q4u4s5 bos taurus
16	448.5	32.5	949	2	Q66X03_MOUSE	Q66x03 mus musculu
17	447	32.4	1003	2	Q66X22_MOUSE	Q66x22 mus musculu
18	436	31.6	1034	1	CIAS1_HUMAN	Q96p20 homo sapien
19	435	31.5	287	2	Q9BY26_HUMAN	Q9by26 homo sapien
20	429	31.1	1061	1	NAL12_HUMAN	P59046 homo sapien
21	424	30.7	989	2	Q6VSG4_MOUSE	Q6vsg4 mus musculu
22	421.5	30.6	994	2	Q86W87_HUMAN	Q86w87 homo sapien
23	420.5	30.5	994	1	NALP4_HUMAN	Q96mn2 homo sapien
24	420.5	30.5	1004	2	Q66X01_MOUSE	Q66x01 mus musculu
25	419	30.4	1033	1	CIAS1_MOUSE	Q8r4b8 mus musculu
26	418.5	30.3	456	1	RINI_PIG	P10775 sus scrofa
27	417.5	30.3	991	2	Q7RTR0_HUMAN	Q7rtr0 homo sapien
28	416	30.2	1062	1	NALP2_HUMAN	Q9nx02 homo sapien
29	416	30.2	1062	2	Q53FL5_HUMAN	Q53fl5 homo sapien
30	410	29.7	456	2	Q5ZIY8_CHICK	Q5ziy8 gallus gall
31	405.5	29.4	713	2	Q95LZ7_MACFA	Q95lz7 macaca fasc
32	401.5	29.1	1043	1	NAL13_HUMAN	Q86w25 homo sapien
33	399.5	29.0	931	2	Q6UQE7_MOUSE	Q6uqe7 mus musculu
34	397	28.8	456	1	RINI_RAT	P29315 rattus norv
35	397	28.8	456	2	Q91VI7_MOUSE	Q91vi7 mus musculu
36	397	28.8	456	2	Q6IRS9_RAT	Q6irs9 rattus norv
37	395	28.6	456	2	Q924P4_MOUSE	Q924p4 mus musculu
38	388.5	28.2	501	2	Q6UTW9_MOUSE	Q6utw9 mus musculu
39	387	28.1	986	2	Q86W27_HUMAN	Q86w27 homo sapien
40	386.5	28.0	1009	2	Q7RTR1_HUMAN	Q7rtr1 homo sapien
41	376.5	27.3	1033	1	NAL11_HUMAN	P59045 homo sapien
42	376.5	27.3	1033	2	Q53ZZ0_HUMAN	Q53zz0 homo sapien
43	375.5	27.2	373	2	Q80YN6_RAT	Q80yn6 rattus norv
44	375.5	27.2	447	2	Q96FD7_HUMAN	Q96fd7 homo sapien
45	375.5	27.2	460	1	RINI_HUMAN	P13489 homo sapien

ALIGNMENTS

RESULT 1
NALP5_HUMAN
ID NALP5_HUMAN STANDARD; PRT; 1200 AA.
AC P59047; Q86W29;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 5 (Mater protein homolog).
GN Name=NALP5; Synonyms=MATER;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RX MEDLINE=21922687; PubMed=11925379; DOI=10.1093/humrep/17.4.903;
RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT "A human homologue of mouse Mater, a maternal effect gene essential
RT for early embryonic development.";
RL Hum. Reprod. 17:903-911(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22451042; PubMed=12563287; DOI=10.1038/nrm1019;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
CC -!- TISSUE SPECIFICITY: Oocyte specific.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 13 LRR (leucine-rich) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC Mater.
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---
DR EMBL; AY054986; AAL15549.1; -; mRNA.
DR EMBL; AY154460; AAO18156.1; -; mRNA.
DR HSSP; P13489; 1A4Y.
DR Ensembl; ENSG00000171487; Homo sapiens.
DR HGNC; HGNC:21269; NALP5.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR Pfam; PF00560; LRR_1; 2.
DR Pfam; PF05729; NACHT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS50824; DAPIN; 1.

```
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Nucleotide-binding; Repeat.
FT   DOMAIN       57    148       DAPIN.
FT   DOMAIN      280    602       NACHT.
FT   REPEAT      704    727       LRR 1.
FT   REPEAT      730    753       LRR 2.
FT   REPEAT      780    803       LRR 3.
FT   REPEAT      809    832       LRR 4.
FT   REPEAT      836    863       LRR 5.
FT   REPEAT      865    892       LRR 6.
FT   REPEAT      893    916       LRR 7.
FT   REPEAT      950    973       LRR 8.
FT   REPEAT      979   1002       LRR 9.
FT   REPEAT     1007   1034       LRR 10.
FT   REPEAT     1036   1059       LRR 11.
FT   REPEAT     1064   1092       LRR 12.
FT   REPEAT     1121   1142       LRR 13.
FT   NP_BIND     286    293       ATP (Potential).
FT   CONFLICT    459    459       I -> M (in Ref. 2).
FT   CONFLICT    506    506       A -> V (in Ref. 2).
FT   CONFLICT    912    912       T -> M (in Ref. 2).
FT   CONFLICT    999    999       S -> F (in Ref. 2).
SQ   SEQUENCE   1200 AA;  134236 MW;  9A070D2A771B28FA CRC64;

  Query Match             100.0%;  Score 1379;  DB 1;  Length 1200;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-110;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 1062

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 1122

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 1182

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||||||||||||||||||
Db       1183 IDGSWHSFDEDDRYWWKN 1200


RESULT 2
Q6VSG0_MOUSE
ID   Q6VSG0_MOUSE PRELIMINARY;      PRT;  1111 AA.
AC   Q6VSG0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MATER protein isoform-B.
GN   Name=Nalp5; Synonyms=Mater;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=A/J;
RX   MEDLINE=22129076; PubMed=12133995;
RA   Roper R.J., Ma R.Z., Biggins J.E., Butterfield R.J., Michael S.D.,
RA   Tung K.S., Doerge R.W., Teuscher C.;
RT   "Interacting quantitative trait loci control loss of peripheral
RT   tolerance and susceptibility to autoimmune ovarian dysgenesis after
RT   day 3 thymectomy in mice.";
RL   J. Immunol. 169:1640-1646(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=A/J;
RA   Ma R.Z., Cheng H., Teuscher C.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY329490; AAQ94613.1; -; mRNA.
DR   MGI; MGI:1345193; Nalp5.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0009887; P:organogenesis; IMP.
DR   InterPro; IPR000767; Disease_resist.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   Pfam; PF00560; LRR_1; 2.
DR   Pfam; PF05729; NACHT; 1.
DR   PRINTS; PR00364; DISEASERSIST.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00368; LRR_RI; 12.
DR   PROSITE; PS50837; NACHT; 1.
SQ   SEQUENCE   1111 AA;  125527 MW;  51EFD4A1A086707E CRC64;

  Query Match             65.8%;  Score 907;  DB 2;  Length 1111;
  Best Local Similarity   65.1%;  Pred. No. 9.6e-70;
  Matches  168;  Conservative   36;  Mismatches   54;  Indels    0;  Gaps    0;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              | ||| ||::||||||||||||| ||| ||: :| | |:||||:|| |::     | ||:
Db        854 LVSALFSNQNLTHLCLSNNSLGTEGVQQLCQFLRNPECALQRLILNHCNIVDDAYGFLAM 913

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              |  |: ||||||:|||| |  :||||| ::||:|:||:|||| | ||  ||| |:|:|:
Db        914 RLANNTKLTHLSLTMNPVGDGAMKLLCEALKEPTCYLQELELVDCQLTQNCCEDLACMIT 973

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ::|||||||| :||||| ||  |||||||  || | |||| || |||:||||||||:|||
Db        974 TTKHLKSLDLGNNALGDKGVITLCEGLKQNNSSLRRLGLGACKLTSNCCEALSLAISCNP 1033

Qy        181 HLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              || |||||:|:||  ||:|||||| || ||| |||||| :|  ::|: ||||: :|| ||
Db       1034 HLNSLNLVKNDFSTSGMLKLCSAFQCPVSNLGIIGLWKQEYYARVRRQLEEVEFVKPHVV 1093

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||| |:: ||||| ||||
Db       1094 IDGDWYASDEDDRNWWKN 1111


RESULT 3
Q6VSG5_MOUSE
ID   Q6VSG5_MOUSE PRELIMINARY;      PRT;  1095 AA.
AC   Q6VSG5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MATER protein isoform-E.
GN   Name=Nalp5; Synonyms=Mater;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SWR/J;
RA   Cheng H., Zhang Y., Huang Z., Ma R.Z.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY329484; AAQ94607.1; -; mRNA.
DR   MGI; MGI:1345193; Nalp5.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0009887; P:organogenesis; IMP.
DR   InterPro; IPR000767; Disease_resist.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:21:18 ; Search time 46.8964 Seconds
(without alignments)
2417.239 Million cell updates/sec

Title: US-10-677-943A-4
Perfect score: 1379
Sequence: 1 LASALVSNRSLTHLCLSNNS.....VVIDGSWHSFDEDDRYWWKN 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1379	100.0	258	5	AAU79524	Aau79524 Human MAT
2	1379	100.0	1200	5	AAU79526	Aau79526 Human MAT
3	1379	100.0	1200	6	AAE31749	Aae31749 Human MAT
4	1379	100.0	1200	9	ADV04681	Adv04681 NALP5, SE
5	1373	99.6	1033	5	ABU65214	Abu65214 Human NOV
6	1373	99.6	1033	8	ADN62079	Adn62079 Human nov
7	1373	99.6	1143	6	ADA45221	Ada45221 Human MAT
8	1373	99.6	1162	5	AAE25053	Aae25053 Human EMB
9	1373	99.6	1162	6	ADA45219	Ada45219 Human MAT
10	1335	96.8	1344	5	AAO15585	Aao15585 Human PYR
11	1335	96.8	1344	7	ADF94745	Adf94745 Human PYR
12	1301	94.3	258	6	AAE31747	Aae31747 Human MAT
13	1293.5	93.8	2312	5	AAO17859	Aao17859 Pyrin dom
14	902	65.4	1111	5	AAU79525	Aau79525 Mouse MAT
15	850	61.6	1111	6	AAE31748	Aae31748 Mouse MAT
16	587.5	42.6	1093	9	ADV44750	Adv44750 Human nuc
17	538.5	39.1	924	7	ADF44705	Adf44705 Human NOV
18	537.5	39.0	924	8	ADO50890	Ado50890 Human NOV
19	524	38.0	1012	8	ADP47734	Adp47734 Human inf
20	477.5	34.6	896	5	AAO15592	Aao15592 Human PYR
21	477.5	34.6	896	7	ADF94759	Adf94759 Human PYR
22	443.5	32.2	218	5	AAU11630	Aau11630 Human Leu
23	436	31.6	330	8	ADP47758	Adp47758 Human CAT
24	436	31.6	629	6	ABR01160	Abr01160 Human gen
25	436	31.6	629	6	ADA98287	Ada98287 Human sec
26	436	31.6	629	7	ADC20452	Adc20452 Human sec
27	436	31.6	629	7	ADF10750	Adf10750 Human sec
28	436	31.6	630	4	AAU01067	Aau01067 Human sec
29	436	31.6	944	8	ADP47756	Adp47756 Human CAT
30	436	31.6	1034	4	AAE07514	Aae07514 Human PYR
31	436	31.6	1034	6	ABU08503	Abu08503 Huma PYRI
32	436	31.6	1034	6	ABU63315	Abu63315 Human pyr
33	436	31.6	1034	7	ADM34353	Adm34353 Human mut
34	436	31.6	1034	7	ADM34355	Adm34355 Human mut
35	436	31.6	1034	7	ADM34351	Adm34351 Human mut
36	436	31.6	1034	7	ADM34354	Adm34354 Human mut
37	436	31.6	1034	7	ADM34319	Adm34319 Human cry
38	436	31.6	1034	7	ADM34352	Adm34352 Human mut
39	436	31.6	1034	8	ADP24014	Adp24014 PRO polyp
40	435	31.5	287	5	AAE21063	Aae21063 rno (upre
41	435	31.5	865	7	ADC31287	Adc31287 Human nov
42	435	31.5	1004	7	ADM04476	Adm04476 Human pro
43	429	31.1	565	5	ABG97475	Abg97475 Human nuc
44	429	31.1	603	5	ADE36457	Ade36457 Human PAN
45	429	31.1	603	8	ADJ19379	Adj19379 Human PAN

ALIGNMENTS

RESULT 1
AAU79524
ID AAU79524 standard; protein; 258 AA.
XX
AC AAU79524;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER peptide encoded by cDNA fragment #2.
XX
KW Human; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; gene therapy; fertility.
XX
OS Homo sapiens.
XX
PN WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI; 2002-454595/48.
DR N-PSDB; ABK48610.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 1; Page 71-72; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,

```
CC   the MATER protein is useful as a contraceptive agent, or for influencing
CC   (either inhibiting or enhancing) fertility and can be used to detect a
CC   predisposition to infertility or reduced fertility, or for presymptomatic
CC   screening of an individual for infertility/reduced fertility. The protein
CC   and polynucleotide are also useful for detecting an excess or deficiency,
CC   or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC   human or a mouse) or for screening for a compound useful in influencing
CC   MATER-mediated fertility. The protein sequence presented is human MATER
CC   which is encoded by cDNA fragment #2, isolated from a human cDNA library
XX
SQ   Sequence 258 AA;

  Query Match            100.0%;  Score 1379;  DB 5;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-126;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180

Qy        181 HLTSLNLVQNNPSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 HLTSLNLVQNNPSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||||||||||||||||||
Db        241 IDGSWHSFDEDDRYWWKN 258


RESULT 2
AAU79526
ID   AAU79526 standard; protein; 1200 AA.
XX
AC   AAU79526;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Human MATER protein.
XX
KW   Human; contraceptive; antiinfertility; MATER;
KW   maternal antigen that embryos require; MATER null phenotype; oocyte;
KW   early embryonic survival; premature ovarian failure; POF;
KW   autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS   Homo sapiens.
XX
PN   WO200232955-A1.
XX
PD   25-APR-2002.
XX
PF   04-APR-2001; 2001WO-US010981.
XX
PR   18-OCT-2000; 2000US-0241510P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nelson LM,  Tong Z;
XX
DR   WPI; 2002-454595/48.
DR   N-PSDB; ABK48628.
XX
PT   New isolated human Maternal Antigen That Embryos Require protein and
PT   polynucleotide, useful in diagnosing or treating fertility or reduced
PT   fertility, or as a contraceptive.
XX
PS   Claim 1; Fig 9; 93pp; English.
XX
CC   The invention discloses an isolated human MATER (Maternal Antigen That
CC   Embryos Require) protein, which can complement a MATER null phenotype in
CC   which zygotes arising from the oocyte do not progress beyond the two-cell
CC   stage. MATER is required for early embryonic survival and abnormal levels
CC   of the protein can lead to premature ovarian failure (POF) and can be
CC   caused by under or over expression of MATER or an autoimmune response to
CC   MATER. MATER is a single-copy maternal effect gene found on chromosome
CC   19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC   diagnosing or treating fertility and reduced fertility. In particular,
CC   the MATER protein is useful as a contraceptive agent, or for influencing
CC   (either inhibiting or enhancing) fertility and can be used to detect a
CC   predisposition to infertility or reduced fertility, or for presymptomatic
CC   screening of an individual for infertility/reduced fertility. The protein
CC   and polynucleotide are also useful for detecting an excess or deficiency,
CC   or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC   human or a mouse) or for screening for a compound useful in influencing
CC   MATER-mediated fertility. The protein sequence presented is the human
CC   MATER. Note: Residues 1-25 are encoded by nucleotides which were deduced,
CC   by the inventors, by comparison to published human genomic DNA sequences,
CC   while the remainder of the sequence was determined by direct cloning of
CC   human ovarian cDNAs
XX
SQ   Sequence 1200 AA;

  Query Match            100.0%;  Score 1379;  DB 5;  Length 1200;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-126;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002

Qy         61 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 1062

Qy        121 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNR 1122

Qy        181 HLTSLNLVQNNPSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 HLTSLNLVQNNPSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVV 1182

Qy        241 IDGSWHSFDEDDRYWWKN 258
              ||||||||||||||||||
Db       1183 IDGSWHSFDEDDRYWWKN 1200


RESULT 3
AAE31749
ID   AAE31749 standard; protein; 1200 AA.
XX
AC   AAE31749;
XX
DT   07-MAR-2003  (first entry)
XX
DE   Human MATER protein.
XX
KW   Human; MATER protein; infertility; fertility; contraceptive agent;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200281492-A1.
XX
PD   17-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009776.
XX
PR   04-APR-2001; 2001WO-US010981.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:38:41 ; Search time 58.9094 Seconds
(without alignments)
2730.707 Million cell updates/sec

Title: US-10-677-943A-2
Perfect score: 2005
Sequence: 1 QAPVTEIMSRPERLLFIIDG.....EVWLPINQNLDLIASSFCLQ 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2005	100.0	682	4	US-10-407-866-96	Sequence 96, Appl
2	2005	100.0	1033	4	US-10-092-900A-348	Sequence 348, App
3	2005	100.0	1143	4	US-10-216-645-4	Sequence 4, Appli
4	2005	100.0	1162	4	US-10-216-645-2	Sequence 2, Appli
5	2005	100.0	1162	4	US-10-416-642-1	Sequence 1, Appli
6	2005	100.0	1344	4	US-10-124-498-6	Sequence 6, Appli
7	2005	100.0	1344	4	US-10-066-521-6	Sequence 6, Appli
8	1997	99.6	1200	4	US-10-399-443-24	Sequence 24, Appl
9	1997	99.6	1200	4	US-10-677-943-24	Sequence 24, Appl
10	1997	99.6	1200	5	US-10-860-761-4	Sequence 4, Appli
11	1940	96.8	385	4	US-10-399-443-2	Sequence 2, Appli
12	1940	96.8	385	4	US-10-677-943-2	Sequence 2, Appli
13	981	48.9	1111	4	US-10-216-645-5	Sequence 5, Appli
14	931	46.4	1111	4	US-10-399-443-6	Sequence 6, Appli
15	931	46.4	1111	4	US-10-677-943-6	Sequence 6, Appli
16	551.5	27.5	612	4	US-10-407-866-88	Sequence 88, Appl
17	543	27.1	628	4	US-10-407-866-98	Sequence 98, Appl
18	543	27.1	732	4	US-10-407-866-72	Sequence 72, Appl
19	543	27.1	920	4	US-10-250-615-13	Sequence 13, Appl
20	543	27.1	924	4	US-10-357-820-52	Sequence 52, Appl
21	543	27.1	1093	5	US-10-794-342-13	Sequence 13, Appl
22	541.5	27.0	674	4	US-10-407-866-89	Sequence 89, Appl
23	541.5	27.0	719	3	US-09-965-621-59	Sequence 59, Appl
24	541.5	27.0	719	4	US-10-407-866-59	Sequence 59, Appl
25	541.5	27.0	719	4	US-10-781-294-59	Sequence 59, Appl
26	541.5	27.0	1032	5	US-10-882-761-33	Sequence 33, Appl
27	541.5	27.0	1034	4	US-10-127-516-5	Sequence 5, Appli
28	541.5	27.0	1034	4	US-10-027-629-5	Sequence 5, Appli
29	541.5	27.0	1034	4	US-10-028-374-18	Sequence 18, Appl
30	541.5	27.0	1034	4	US-10-132-967-5	Sequence 5, Appli
31	541.5	27.0	1034	4	US-10-183-770-18	Sequence 18, Appl
32	541.5	27.0	1034	4	US-10-264-958B-22	Sequence 22, Appl
33	541.5	27.0	1034	6	US-11-107-572-18	Sequence 18, Appl
34	534.5	26.7	599	4	US-10-407-866-99	Sequence 99, Appl
35	534.5	26.7	635	4	US-10-407-866-90	Sequence 90, Appl
36	534.5	26.7	858	3	US-09-848-035-8	Sequence 8, Appli
37	534.5	26.7	858	3	US-09-986-224-8	Sequence 8, Appli
38	534.5	26.7	858	5	US-10-116-422-8	Sequence 8, Appli
39	534.5	26.7	919	4	US-10-094-749-2718	Sequence 2718, Ap
40	534.5	26.7	919	4	US-10-161-493-4	Sequence 4, Appli
41	534.5	26.7	919	4	US-10-398-037-7	Sequence 7, Appli
42	534.5	26.7	994	3	US-09-965-621-16	Sequence 16, Appl
43	534.5	26.7	994	4	US-10-124-498-24	Sequence 24, Appl
44	534.5	26.7	994	4	US-10-066-521-24	Sequence 24, Appl
45	534.5	26.7	994	4	US-10-407-866-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-10-407-866-96
; Sequence 96, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: 66654-10(LJ 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 682
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-96

Query Match 100.0%; Score 2005; DB 4; Length 682;
Best Local Similarity 100.0%; Pred. No. 6.9e-193;
Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 352

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 412

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 472

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 532

Qy        241 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 592

Qy        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 652
```

```
Qy        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
              |||||||||||||||||||||||||
Db        653 LNSFQEVWLPINQNLDLIASSFCLQ 677
```

RESULT 2
US-10-092-900A-348
; Sequence 348, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev,, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 348
; LENGTH: 1033
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-348

```
  Query Match           100.0%;  Score 2005;  DB 4;  Length 1033;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-192;
  Matches  385;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 236

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 296

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 356

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 416

Qy        241 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 476

Qy        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 536

Qy        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
              |||||||||||||||||||||||||
Db        537 LNSFQEVWLPINQNLDLIASSFCLQ 561
```

RESULT 3
US-10-216-645-4
; Sequence 4, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA
; APPLICANT: PETERS-KOTTIG, MICHAELE
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN MATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-645-4

```
  Query Match           100.0%;  Score 2005;  DB 4;  Length 1143;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-192;
  Matches  385;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 346

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 406

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 466

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:39:46 ; Search time 2.50678 Seconds
(without alignments)
465.507 Million cell updates/sec

Title: US-10-677-943A-2
Perfect score: 2005
Sequence: 1 QAPVTEIMSRPERLLFIIDG.....EVWLPINQNLDLIASSFCLQ 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Score	% Query Match	Length	DB	ID	Description
1	457.5	22.8	1062	7	US-11-137-465-43	Sequence 43, Appl
2	433.5	21.6	1049	7	US-11-137-465-42	Sequence 42, Appl
3	88	4.4	499	1	US-10-821-234-1106	Sequence 1106, Ap
4	81.5	4.1	568	1	US-10-793-626-2482	Sequence 2482, Ap
5	81.5	4.1	1145	1	US-10-793-626-1432	Sequence 1432, Ap
6	79.5	4.0	1217	7	US-11-074-176-252	Sequence 252, App
7	79.5	4.0	3056	7	US-11-109-156-20	Sequence 20, Appl
8	77.5	3.9	622	7	US-11-021-441-35	Sequence 35, Appl
9	74.5	3.7	518	1	US-10-821-234-1165	Sequence 1165, Ap
10	72	3.6	265	1	US-10-392-234A-54	Sequence 54, Appl
11	71.5	3.6	745	1	US-10-793-626-1742	Sequence 1742, Ap
12	71.5	3.6	1150	1	US-10-821-234-1083	Sequence 1083, Ap
13	70	3.5	512	7	US-11-012-762-64	Sequence 64, Appl
14	70	3.5	621	1	US-10-821-234-1376	Sequence 1376, Ap
15	70	3.5	882	7	US-11-012-762-34	Sequence 34, Appl
16	70	3.5	1451	7	US-11-046-346-1	Sequence 1, Appli
17	69	3.4	667	1	US-10-793-626-198	Sequence 198, App
18	67.5	3.4	754	1	US-10-793-626-1296	Sequence 1296, Ap
19	67.5	3.4	775	1	US-10-131-826A-120	Sequence 120, App
20	67.5	3.4	932	7	US-11-017-550-65	Sequence 65, Appl
21	67	3.3	750	1	US-10-689-742-86	Sequence 86, Appl
22	66.5	3.3	199	1	US-10-793-626-186	Sequence 186, App
23	66.5	3.3	309	1	US-10-858-730-213	Sequence 213, App
24	66.5	3.3	357	1	US-10-793-626-410	Sequence 410, App
25	66	3.3	1087	7	US-11-102-978-2	Sequence 2, Appli
26	65.5	3.3	293	1	US-10-485-517-203	Sequence 203, App
27	65	3.2	317	7	US-11-051-568-11	Sequence 11, Appl
28	65	3.2	408	7	US-11-051-568-15	Sequence 15, Appl
29	65	3.2	430	7	US-11-051-568-25	Sequence 25, Appl
30	65	3.2	431	1	US-10-816-768-39	Sequence 39, Appl
31	65	3.2	431	7	US-11-051-568-2	Sequence 2, Appli
32	65	3.2	484	7	US-11-051-568-13	Sequence 13, Appl
33	64.5	3.2	396	7	US-11-074-176-346	Sequence 346, App
34	64.5	3.2	405	7	US-11-074-176-164	Sequence 164, App
35	64.5	3.2	451	1	US-10-131-826A-126	Sequence 126, App
36	64.5	3.2	495	7	US-11-037-829A-7	Sequence 7, Appli
37	64.5	3.2	683	1	US-10-821-234-1314	Sequence 1314, Ap
38	64.5	3.2	683	7	US-11-081-566-2	Sequence 2, Appli
39	64.5	3.2	696	1	US-10-131-826A-354	Sequence 354, App
40	64.5	3.2	999	1	US-10-821-234-1251	Sequence 1251, Ap
41	64.5	3.2	1062	1	US-10-821-234-1079	Sequence 1079, Ap
42	64	3.2	333	7	US-11-004-789-2	Sequence 2, Appli
43	64	3.2	665	1	US-10-793-626-1232	Sequence 1232, Ap
44	64	3.2	1043	1	US-10-392-234A-34	Sequence 34, Appl
45	63.5	3.2	511	1	US-10-131-826A-122	Sequence 122, App

ALIGNMENTS

RESULT 1
US-11-137-465-43
; Sequence 43, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 1062
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-43

Query Match 22.8%; Score 457.5; DB 7; Length 1062;
Best Local Similarity 31.8%; Pred. No. 5.9e-39;
Matches 127; Conservative 69; Mismatches 165; Indels 39; Gaps 10;

```
Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGS---VLNNDTKLCKDWAEKQPPFTLIRSLLRKVLL 57
              |  :  |:::  ::||:|||||:||:    |  |  :| || :|:|   |: ||| :|:|
Db        270 QDDIPHILAQARKILFVIDGFDELGAAPGALIED--ICGDWEKKKPVPVLLGSLLNRVML 327

Qy         58 PESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMN 117
              |:: |:|| |     |:   | |: | |   | |    |  | :    : :
Db        328 PKAALLVTTRPRALRDLRILAEEPIYIRVEGFLEEDRRAYFLRHFGDEDQAMRAFELMRS 387

Qy        118 NRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCL 177
              |  |      ||| ::|  |:||   ||   |  | |||   |:  :  |:|
Db        388 NAALFQLGSAPAVCWIVCTTLKLQMEKGEDPVPTCLTRTGLFLRFLCSRF-PQGA----- 441
```

```
Qy        178 NLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYY 237
Db        442 --QLRGALRTLSLLAAQGLWAQTSVLHREDLERLGVQESDLRLFLDGDILRQDRVSKGCY 499

Qy        238 TFFHLSLQDFCAALYYVLEGLEIEPALCPLY-----------VEKTKRSMELKQAGFHIH 286
Db        500 SFIHLSFQQFLTALFYTLEKEEEDRDGHTWDIGDVQKLLSGVERL-RNPDLIQAGY--- 555

Qy        287 SLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHW-VSLLGQQPNATTPGDTLDAF 345
Db        556 ------YSFGLANEKRAKELEATFGCRMSPDIKQELLRCDISCKGGHSTVT---DLQELL 606

Qy        346 HCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQ 385
Db        607 GCLYESQEEELVKEVMAQFKEISLHLNA-VDVVPSSFCVK 645


RESULT 2
US-11-137-465-42
; Sequence 42, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
;  APPLICANT: Agarwal, Pankaj
;  APPLICANT:  Murdoch, Paul R.
;  APPLICANT:  Rizvi, Safia, K.
;  APPLICANT:  Smith, Randall, F.
;  APPLICANT:  Xiang, Zhaoying
;  APPLICANT:  Kabnick, Karen
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP50018
;  CURRENT APPLICATION NUMBER: US/11/137,465
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US/10/239,663
;  PRIOR FILING DATE: 2002-09-24
;  PRIOR APPLICATION NUMBER: PCT/US01/09226
;  PRIOR FILING DATE: 2001-03-22
;  PRIOR APPLICATION NUMBER: 60/192,158
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,668
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: 60/200,166
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 42
;   LENGTH: 1049
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-137-465-42

  Query Match           21.6%;  Score 433.5;  DB 7;  Length 1049;
  Best Local Similarity   29.5%;  Pred. No. 1.6e-36;
  Matches  118;  Conservative   74;  Mismatches  171;  Indels   37;  Gaps    8;

Qy          1 QAPVTEIMSRPERLLFIIDGFDDL----GSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVL 56
Db        261 QDDIPSILAQAQRILFVVDGLDELKVPPGALIQD---ICGDWEKKKPVPVLLGSLLKRKM 317

Qy         57 LPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIM 116
Db        318 LPRAALLVTTRPRALRDLQLLAQQPIYVRVEGFLEEDRRAYFLRHFGDEDQAMRAFELMR 377

Qy        117 NNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRC 176
Db        378 SNAALFQLGSAPAVCWIVCTTLKLQMEKGEDPVPTCLTRTGLFLRFLCSRF-PQGA---- 432

Qy        177 LNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEY 236
Db        433 ---QLRGALRTLSLLAAQGLWAQMSVFHREDLERLGVQESDLRLFLDGDILRQDRVSKGC 489

Qy        237 YTFFHLSLQDFCAALYYVLE----------GLEIEPALCPLYVEKTKRSMELKQAGFHIH 286
Db        490 YSFIHLSFQQFLTALFYALEKEEGEDRDGHAWDIGDVQKLLSGEERLKNPDLIQVG---- 545

Qy        287 SLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHW-VSLLGQQPNATTPGDTLDAF 345
Db        546 -----HFLFGLANEKRAKELEATFGCRMSPDIKQELLQCKAHLHANKPLSVT--DLKEVL 598

Qy        346 HCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQ 385
Db        599 GCLYESQEEELAKVVVAPFKEISIHLTNTSEVMHCSFSLK 638


RESULT 3
US-10-821-234-1106
; Sequence 1106, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1106
;   LENGTH: 499
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-821-234-1106

  Query Match           4.4%;  Score 88;  DB 1;  Length 499;
  Best Local Similarity   19.1%;  Pred. No. 0.1;
  Matches   52;  Conservative   45;  Mismatches  115;  Indels   60;  Gaps    8;

Qy         45 FTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIG 104
Db        193 FTILPAILQSAALPFC----------------PESPRFLLINR-KBEENAKQILQRLWG 234

Qy        105 EHQKTQGLRAIMNNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVF 164
Db        235 TQDVSQDIQEMKDESARMSQEKQVTVLELFRVSSYRQPIIISIVLQLSQQLSGINAVFYY 294

Qy        165 -------------HQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMV 210
Db        295 STGIFKDAGVQEPIYATIGAGVVNTIFTVVSLFLVERAGRRTLHMI-------------- 340

Qy        211 QGLGESEL-RALFHMNILLPDSHCEEYYTFFHLSLQDFC-AALYYVLEGLEIEPALCPLY 268
Db        341 -GLGGMAFCSTLMTVSLLLKDN---------YNGMSFVCIGAILVFVAFFEIGPGPIPWF 390

Qy        269 VEKTKRSMELKQAGFHIH--SLWMKRFLFGLV 298
Db        391 IVAELFSQGPRPAAMAVAGCSNWTSNFLVGLL 422


RESULT 4
US-10-793-626-2482
; Sequence 2482, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/10/793,626
;  CURRENT FILING DATE:  2004-03-04
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:26:25 ; Search time 18.592 Seconds
(without alignments)
1712.037 Million cell updates/sec

Title: US-10-677-943A-2
Perfect score: 2005
Sequence: 1 QAPVTEIMSRPERLLFIIDG.....EVWLPINQNLDLIASSFCLQ 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	541.5	27.0	719	2	US-10-781-294-59	Sequence 59, Appl
2	541.5	27.0	1034	2	US-10-183-770A-18	Sequence 18, Appl
3	534.5	26.7	994	2	US-10-781-294-16	Sequence 16, Appl
4	534.5	26.7	1035	2	US-10-781-294-24	Sequence 24, Appl
5	484.5	24.2	1399	2	US-09-388-221B-4	Sequence 4, Appli
6	484.5	24.2	1424	2	US-09-388-221B-12	Sequence 12, Appl
7	484.5	24.2	1429	2	US-10-029-347-3	Sequence 3, Appli
8	484.5	24.2	1429	2	US-10-029-347-26	Sequence 26, Appl
9	484.5	24.2	1429	2	US-10-183-770A-3	Sequence 3, Appli
10	484.5	24.2	1429	2	US-10-183-770A-15	Sequence 15, Appl
11	484.5	24.2	1429	2	US-09-996-617-2	Sequence 2, Appli
12	484.5	24.2	1443	2	US-09-388-221B-6	Sequence 6, Appli
13	484.5	24.2	1454	2	US-09-388-221B-10	Sequence 10, Appl
14	484.5	24.2	1473	2	US-09-388-221B-2	Sequence 2, Appli
15	454.5	22.7	1033	2	US-10-029-347-4	Sequence 4, Appli
16	449	22.4	449	2	US-10-183-770A-2	Sequence 2, Appli
17	400.5	20.0	947	2	US-10-781-294-18	Sequence 18, Appl
18	398	19.9	681	2	US-10-781-294-20	Sequence 20, Appl
19	385.5	19.2	319	2	US-10-781-294-37	Sequence 37, Appl
20	384.5	19.2	625	2	US-10-029-347-2	Sequence 2, Appli
21	358	17.9	514	2	US-10-029-347-25	Sequence 25, Appl
22	358	17.9	514	2	US-10-183-770A-14	Sequence 14, Appl
23	353	17.6	270	2	US-10-781-294-38	Sequence 38, Appl
24	316.5	15.8	603	2	US-10-781-294-64	Sequence 64, Appl
25	309.5	15.4	504	2	US-10-781-294-22	Sequence 22, Appl
26	269	13.4	317	2	US-10-781-294-60	Sequence 60, Appl
27	235.5	11.7	209	2	US-09-388-221B-19	Sequence 19, Appl
28	184.5	9.2	190	2	US-10-781-294-63	Sequence 63, Appl
29	177	8.8	1114	2	US-09-949-002-550	Sequence 550, App
30	171	8.5	830	2	US-09-064-199-19	Sequence 19, Appl
31	171	8.5	1106	2	US-09-064-199-17	Sequence 17, Appl
32	171	8.5	1130	1	US-08-519-547A-6	Sequence 6, Appli
33	171	8.5	1130	2	US-09-064-199-18	Sequence 18, Appl
34	171	8.5	1130	2	US-09-949-002-295	Sequence 295, App
35	171	8.5	1207	2	US-09-064-199-16	Sequence 16, Appl
36	154	7.7	181	2	US-10-781-294-62	Sequence 62, Appl
37	140.5	7.0	953	2	US-09-245-281-43	Sequence 43, Appl
38	140.5	7.0	953	2	US-09-207-359B-43	Sequence 43, Appl
39	140.5	7.0	953	2	US-09-340-620A-43	Sequence 43, Appl
40	140.5	7.0	953	2	US-09-865-364-43	Sequence 43, Appl
41	140.5	7.0	953	2	US-09-728-721-43	Sequence 43, Appl
42	136	6.8	953	2	US-09-099-041A-8	Sequence 8, Appli
43	136	6.8	953	2	US-09-245-281-8	Sequence 8, Appli
44	136	6.8	953	2	US-09-207-359B-8	Sequence 8, Appli
45	136	6.8	953	2	US-09-340-620A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-781-294-59
; Sequence 59, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58,
; LOCATION: 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71
; OTHER INFORMATION: Xaa = Any Amino Acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86,
; LOCATION: 87, 88, 89, 90, 91, 92, 93, 94, 95
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-781-294-59

Query Match 27.0%; Score 541.5; DB 2; Length 719;
Best Local Similarity 29.7%; Pred. No. 2.1e-52;
Matches 120; Conservative 93; Mismatches 152; Indels 39; Gaps 6;

```
Qy          3 PVTEIMSRPERLLFIIDGFDDL-GSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESF 61
              |: :|: :| |:||::||||:| |:   :   || || : :    |: ||:|| ||||:
Db        286 PIHKIVRKPSRILFLMDGPDELQGAFDEHIGPLCTDWQKAERGDILLSSLIRKKLLPEAS 345

Qy         62 LIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNREL 121
              |::| | |  |||:  :  ||:: : | |  :|     :    | |       | | |
Db        346 LLITTRPVALEKLQHLLDHPRHVEILGFSEAKRKEYFFKYFSDEAQARAAFSLIQENEVL 405
```

```
Qy        122 LDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRR---CLN 178
Db        406 FTMCFIPLVCWIVCTGLKQQMESGKSLAQTSKTTTAVYVFFLSSLLQPRGGSQEHGLCAH 465

Qy        179 LEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYT 238
Db        466 ------LWGLCSLAADGIWNQKILFEESDLRNHGLQKADVSAFLRMNLFQKEVDCEKFYS 519

Qy        239 FFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRS------MELKQAGFHI------- 285
Db        520 FIHMTFQEFFAAMYYLLE-----------EEKEGRTNVPGSRLKLPSRDVTVLLENYGK 567

Qy        286 ----HSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDT 341
Db        568 FEKGYLIFVVRFLFGLVNQERTSYLEKKLSCKISQQIRLELLKWIEVKAKAKKLQIQPSQ 627

Qy        342 LDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQ 385
Db        628 LELFYCLYEMQEEDFVQRAMDYFPKIEINLSTRMDHMVSSFCIE 671
```

RESULT 2
US-10-183-770A-18
; Sequence 18, Application US/10183770A
; Patent No. 6949363
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRBM1
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770A
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-183-770A-18

Query Match 27.0%; Score 541.5; DB 2; Length 1034;
Best Local Similarity 29.7%; Pred. No. 3.8e-52;
Matches 120; Conservative 93; Mismatches 152; Indels 39; Gaps 6;

```
Qy          3 PVTEIMSRPERLLFIIDGFDDL-GSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESF 61
Db        284 PIHKIVRKPSRILFLMDGFDELQGAFDEHIGPLCTDWQKAERGDILLSSLIRKKLLPEAS 343

Qy         62 LIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNREL 121
Db        344 LLITTRPVALEKLQHLLDHPRHVEILGFSEAKRKEYFFKYFSDEAQARAAFSLIQENEVL 403

Qy        122 LDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRR---CLN 178
Db        404 FTMCFIPLVCWIVCTGLKQQMESGKSLAQTSKTTTAVYVFFLSSLLQPRGGSQEHGLCAH 463

Qy        179 LEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYT 238
Db        464 ------LWGLCSLAADGIWNQKILFEESDLRNHGLQKADVSAFLRMNLFQKEVDCEKFYS 517

Qy        239 FFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRS------MELKQAGFHI------- 285
Db        518 FIHMTFQEFFAAMYYLLE-----------EEKEGRTNVPGSRLKLPSRDVTVLLENYGK 565

Qy        286 ----HSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDT 341
Db        566 FEKGYLIFVVRFLFGLVNQERTSYLEKKLSCKISQQIRLELLKWIEVKAKAKKLQIQPSQ 625

Qy        342 LDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQ 385
Db        626 LELFYCLYEMQEEDFVQRAMDYFPKIEINLSTRMDHMVSSFCIE 669
```

RESULT 3
US-10-781-294-16
; Sequence 16, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-16

Query Match 26.7%; Score 534.5; DB 2; Length 994;
Best Local Similarity 31.8%; Pred. No. 2.2e-51;
Matches 125; Conservative 82; Mismatches 165; Indels 21; Gaps 4;

```
Qy          2 APVTEIMSRPERLLFIIDGFDDLGSVLNN-DTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
Db        214 APITEIVSQPERLLFVIDSFEELQGGLNEPDSDLCGDLMEKRPVQVLLSSLLRKKMLPEA 273

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
Db        274 SLLIAIKPVCPKELRDQVTISEIYQPRGFNESDRLVYFCCFFKDPKRAMEAFNLVRESEQ 333

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
Db        334 LFSICQIPLLCWILCTSLKQEMQKGKDLALTCQSTTSVYSSFVFNLFTPEGAEGPTPQTQ 393

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
Db        394 HQ--LKALCSLAAEGMWTDTFEFCEDDLRRNGVVDADIPALLGTKILLKYGERESSYVFL 451

Qy        241 HLSLQDFCAALYYVLEGLEIEP--------ALCPLYVEKTKRSMELKQAGFHIHSLWMKR 292
Db        452 HVCIQEFCAALFYLLKSHLDHPHPAVRCVQELLVANFEKARRA----------HWIFLGC 501

Qy        293 FLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQ 352
Db        502 FLTGLLNKKEQEKLDAFFGFQLSQEIKQQIHQCLKSLGERGNPQGQVDSLAIFYCLFEMQ 561

Qy        353 DKEFVRLALNSFQEVWLPINQNLDLIASSFCLQ 385
Db        562 DPAFVKQAVNLLQEANFHIIDNVDLVVSAYCLK 594
```

RESULT 4
US-10-781-294-24
; Sequence 24, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:22:06 ; Search time 75.4124 Seconds
(without alignments)
3601.908 Million cell updates/sec

Title: US-10-677-943A-2
Perfect score: 2005
Sequence: 1 QAPVTEIMSRPERLLFIIDG.....EVWLPINQNLDLIASSFCLQ 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1997	99.6	1200	1	NALP5_HUMAN	P59047 homo sapien
2	988	49.3	727	2	Q6VSG2_MOUSE	Q6vsg2 mus musculu
3	988	49.3	1085	2	Q6VSG1_MOUSE	Q6vsg1 mus musculu
4	988	49.3	1095	2	Q6VSG3_MOUSE	Q6vsg3 mus musculu
5	988	49.3	1111	2	Q6XZF2_MOUSE	Q6xzf2 mus musculu
6	981	48.9	727	2	Q6VSF9_MOUSE	Q6vsf9 mus musculu
7	981	48.9	989	2	Q6VSG4_MOUSE	Q6vsg4 mus musculu
8	981	48.9	1095	2	Q6VSG5_MOUSE	Q6vsg5 mus musculu
9	981	48.9	1111	1	NALP5_MOUSE	Q9r1m5 mus musculu
10	981	48.9	1111	2	Q53ZD5_MOUSE	Q53zd5 mus musculu
11	981	48.9	1111	2	Q6VSG0_MOUSE	Q6vsg0 mus musculu
12	981	48.9	1111	2	Q7TPU9_MOUSE	Q7tpu9 mus musculu
13	953.5	47.6	1098	2	Q647I9_BOVIN	Q647i9 bos taurus
14	559.5	27.9	506	2	Q68U58_9PRIM	Q68u58 cercopithec
15	555.5	27.7	506	2	Q68U60_THEGE	Q68u60 theropithec
16	553.5	27.6	506	2	Q68U59_MACSY	Q68u59 macaca sylv
17	553	27.6	475	2	Q68U48_CHEME	Q68u48 cheirogaleu
18	551.5	27.5	499	2	Q68U55_AOTAZ	Q68u55 aotus azara
19	551.5	27.5	1033	1	CIAS1_MOUSE	Q8r4b8 mus musculu
20	551	27.5	500	2	Q68U51_CALJA	Q68u51 callithrix
21	549.5	27.4	506	2	Q68U57_COLGU	Q68u57 colobus gue
22	548	27.3	500	2	Q68U50_CEBPY	Q68u50 cebuella py
23	546.5	27.3	506	2	Q68U61_9PRIM	Q68u61 gorilla gor
24	545	27.2	500	2	Q68U56_ALOBE	Q68u56 alouatta be
25	543.5	27.1	437	2	Q68U49_TARBA	Q68u49 tarsius ban
26	543	27.1	1093	1	NAL14_HUMAN	Q86w24 homo sapien
27	541.5	27.0	1034	1	CIAS1_HUMAN	Q96p20 homo sapien
28	540	26.9	500	2	Q68U52_SAGMI	Q68u52 saguinus mi
29	534.5	26.7	994	1	NALP4_HUMAN	Q96mn2 homo sapien
30	534.5	26.7	1061	1	NAL12_HUMAN	P59046 homo sapien
31	533.5	26.6	994	2	Q86W87_HUMAN	Q86w87 homo sapien
32	530	26.4	500	2	Q68U54_SAISC	Q68u54 saimiri sci
33	517	25.8	500	2	Q68U53_CEBAP	Q68u53 cebus apell
34	515	25.7	413	2	Q6JEK9_LOXAF	Q6jek9 loxodonta a
35	499	24.9	413	2	Q6JEK7_DASNO	Q6jek7 dasypus nov
36	488	24.3	986	2	Q66X14_MOUSE	Q66x14 mus musculu
37	488	24.3	993	2	Q6B966_MOUSE	Q6b966 mus musculu
38	484.5	24.2	1375	2	Q86UB5_HUMAN	Q86ub5 homo sapien
39	484.5	24.2	1473	1	NALP1_HUMAN	Q9c000 homo sapien
40	478	23.8	719	2	Q6UQE6_MOUSE	Q6uqe6 mus musculu
41	477	23.8	1043	1	NAL13_HUMAN	Q86w25 homo sapien
42	469.5	23.4	370	2	Q6JEK8_ERIEU	Q6jek8 erinaceus e
43	457.5	22.8	761	2	Q59G09_HUMAN	Q59g09 homo sapien
44	457.5	22.8	846	2	Q8IXT0_HUMAN	Q8ixt0 homo sapien
45	457.5	22.8	1062	1	NALP2_HUMAN	Q9nx02 homo sapien

ALIGNMENTS

RESULT 1
NALP5_HUMAN
ID NALP5_HUMAN STANDARD; PRT; 1200 AA.
AC P59047; Q86W29;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 5 (Mater protein homolog).
GN Name=NALP5; Synonyms=MATER;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RX MEDLINE=21922687; PubMed=11925379; DOI=10.1093/humrep/17.4.903;
RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT "A human homologue of mouse Mater, a maternal effect gene essential
RT for early embryonic development.";
RL Hum. Reprod. 17:903-911(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22451042; PubMed=12563287; DOI=10.1038/nrm1019;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
CC -!- TISSUE SPECIFICITY: Oocyte specific.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 13 LRR (leucine-rich) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC Mater.
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---
DR EMBL; AY054986; AAL15549.1; -; mRNA.
DR EMBL; AY154460; AAO18156.1; -; mRNA.
DR HSSP; P13489; 1A4Y.
DR Ensembl; ENSG00000171487; Homo sapiens.
DR HGNC; HGNC:21269; NALP5.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR Pfam; PF00560; LRR_1; 2.
DR Pfam; PF05729; NACHT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS50824; DAPIN; 1.

```
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Nucleotide-binding; Repeat.
FT   DOMAIN       57    148       DAPIN.
FT   DOMAIN      280    602       NACHT.
FT   REPEAT      704    727       LRR 1.
FT   REPEAT      730    753       LRR 2.
FT   REPEAT      780    803       LRR 3.
FT   REPEAT      809    832       LRR 4.
FT   REPEAT      836    863       LRR 5.
FT   REPEAT      865    892       LRR 6.
FT   REPEAT      893    916       LRR 7.
FT   REPEAT      950    973       LRR 8.
FT   REPEAT      979   1002       LRR 9.
FT   REPEAT     1007   1034       LRR 10.
FT   REPEAT     1036   1059       LRR 11.
FT   REPEAT     1064   1092       LRR 12.
FT   REPEAT     1121   1142       LRR 13.
FT   NP_BIND     286    293       ATP (Potential).
FT   CONFLICT    459    459       I -> M (in Ref. 2).
FT   CONFLICT    506    506       A -> V (in Ref. 2).
FT   CONFLICT    912    912       T -> M (in Ref. 2).
FT   CONFLICT    999    999       S -> F (in Ref. 2).
SQ   SEQUENCE   1200 AA;  134236 MW;  9A070D2A771B28FA CRC64;

  Query Match             99.6%;  Score 1997;  DB 1;  Length 1200;
  Best Local Similarity   99.5%;  Pred. No. 5.1e-163;
  Matches  383;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 403

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        404 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRE 463

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        464 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLE 523

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 583

Qy        241 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 643

Qy        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 703

Qy        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
              |||||||||||||||||||||||||
Db        704 LNSFQEVWLPINQNLDLIASSFCLQ 728


RESULT 2
Q6VSG2_MOUSE
ID   Q6VSG2_MOUSE PRELIMINARY;      PRT;   727 AA.
AC   Q6VSG2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MATER protein isoform-F.
GN   Name=Nalp5; Synonyms=Mater;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CBA/J;
RA   Cheng H., Zhang Y., Huang Z., Ma R.Z.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY329488; AAQ94611.1; -; mRNA.
DR   MGI; MGI:1345193; Nalp5.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0009887; P:organogenesis; IMP.
DR   InterPro; IPR000767; Disease_resist.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   Pfam; PF05729; NACHT; 1.
DR   PRINTS; PR00364; DISEASERSIST.
DR   PROSITE; PS50837; NACHT; 1.
SQ   SEQUENCE   727 AA;  82942 MW;  84CF51ECC04B5BB9 CRC64;

  Query Match             49.3%;  Score 988;  DB 2;  Length 727;
  Best Local Similarity   54.7%;  Pred. No. 3.4e-76;
  Matches  211;  Conservative   55;  Mismatches  114;  Indels    6;  Gaps    5;

Qy          3 PVTEIMSRPERLLFIIDGFDDLGSVL-NNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESF 61
              |||:|||:||||||:||| ||: ||| ::|  | :|| ::|| : |: |||||| |||:||
Db        256 PVTKIMSQPERLLFVIDGLDDMDSVLQHDDMTLSRDWKDEQPIYILMYSLLRKALLPQSF 315

Qy         62 LIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNREL 121
              ||:| |: | ||||| |||| |:|| |:|  :|  |:||     |  : |   ::: | :|
Db        316 LIITTRNTGLEKLKSMVVSPLYILVEGLSASRRSQLVLENISNESDRIQVFHSLIENHQL 375

Qy        122 LDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLEE 181
              |||| |:| ||:| |||||  :|:      ||||||:|  |||||| :   :  |: ||
Db        376 FDQCQAPSVCSLVCEALQLQKKLGKRCTLPCQTLTGLYATLVFHQLTLKRPSQSALSQEE 435

Qy        182 RVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSH-CEEYYTFF 240
              :: |   | || ||||  :||| |||     | |||: |||||||||   |  |: | |
Db        436 QITLVGLCMMAAEGVWTMRSVFYDDDLKNYSLKESEILALFHMNILLQVGHNSEQCYVFS 495

Qy        241 HLSLQDFCAALYYVLEGL-EIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVS 299
              ||||||| |||||||||| |      |  ::|  :  ||:|:       | ||||||||::
Db        496 HLSLQDFFAALYYVLEGLEEWNQHFC--FIENQRSIMEVKRTD-DTRLLGMKRFLFGLMN 552

Qy        300 EDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRL 359
              :|: : ||||   ||   |:||| |||||: || | |:| |||||:||||:|||:|||
Db        553 KDILKTLEVLFEYPVIPTVEQKLQHWVSLIAQQVNGTSPMDTLDAFYCLFESQDEEFVGG 612

Qy        360 ALNSFQEVWLPINQNLDLIASSFCLQ 385
              ||  |||||| ||| :||  ||:||:
Db        613 ALKRFQEVWLLINQKMDLKVSSYCLK 638


RESULT 3
Q6VSG1_MOUSE
ID   Q6VSG1_MOUSE PRELIMINARY;      PRT;  1085 AA.
AC   Q6VSG1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MATER protein isoform-G.
GN   Name=Nalp5; Synonyms=Mater;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CBA/J;
RA   Cheng H., Zhang Y., Huang Z., Ma R.Z.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY329489; AAQ94612.1; -; mRNA.
DR   HSSP; P13489; 1A4Y.
DR   MGI; MGI:1345193; Nalp5.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0009887; P:organogenesis; IMP.
DR   InterPro; IPR000767; Disease_resist.
DR   InterPro; IPR001611; LRR.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:22:26 ; Search time 14.8318 Seconds
(without alignments)
2497.570 Million cell updates/sec

Title: US-10-677-943A-2
Perfect score: 2005
Sequence: 1 QAPVTEIMSRPERLLFIIDG.....EVWLPINQNLDLIASSFCLQ 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	981	48.9	1111	2	A59000	mater protein [imp
2	484.5	24.2	1192	2	T17255	hypothetical prote
3	223.5	11.1	483	2	S27880	Nasopressin recept
4	171	8.5	1130	2	A48843	MHC class II trans
5	133.5	6.7	1075	2	T31668	hypothetical prote
6	123	6.1	1004	2	T31665	hypothetical prote
7	113	5.6	1526	2	AC2239	WD-40 repeat prote
8	102.5	5.1	770	2	F65028	Penicillin-binding
9	101.5	5.1	727	2	E85896	probable peptidogl
10	101.5	5.1	770	2	A91052	penicillin binding
11	98.5	4.9	889	2	C86257	resistance to Pseu
12	93	4.6	247	2	S66972	probable membrane
13	92	4.6	437	2	S22307	site-specific DNA-
14	92	4.6	439	2	F75306	probable sun prote
15	91.5	4.6	840	2	T21333	hypothetical prote
16	90.5	4.5	337	2	D64638	hypothetical prote
17	90.5	4.5	620	2	A53731	translation initia
18	90.5	4.5	1273	2	T34558	hypothetical prote
19	90	4.5	863	2	AH2296	hypothetical prote
20	90	4.5	1562	2	S53069	probable membrane
21	89.5	4.5	337	2	C71875	hypothetical prote
22	89.5	4.5	451	2	A96718	hypothetical prote
23	89.5	4.5	763	2	G97026	superfamily I DNA
24	89	4.4	326	2	S77304	beta transducin-li
25	88.5	4.4	671	2	AD1666	DNA ligase homolog
26	88	4.4	496	2	A31986	glucose transporte
27	87.5	4.4	323	2	H82082	RNA-directed DNA p
28	87.5	4.4	416	2	G97585	reca protein (reco
29	87.5	4.4	608	2	B87282	DNA polymerase III
30	87.5	4.4	667	2	C87236	probable zinc meta
31	87.5	4.4	668	1	F69794	DNA ligase (NAD) (
32	87	4.3	1574	2	G70466	DNA-directed RNA p
33	86.5	4.3	433	2	B64603	anaerobic glycerol
34	86.5	4.3	450	2	H71135	probable adenylosu
35	86.5	4.3	491	2	T16691	hypothetical prote
36	86.5	4.3	771	2	AB0823	penicillin-binding
37	86.5	4.3	3795	2	T00831	hypothetical prote
38	86	4.3	588	2	AG1148	a fusion of two ty
39	86	4.3	589	2	T40252	hypothetical prote
40	86	4.3	632	2	S58152	hypothetical prote
41	86	4.3	2701	2	S17796	inositol-trisphosp
42	85.5	4.3	600	2	T01062	hypothetical prote
43	85.5	4.3	1517	2	B81393	DNA-directed RNA p
44	85	4.2	423	2	G96033	conserved hypothet
45	84.5	4.2	352	2	G86660	adenosine deaminas

ALIGNMENTS

RESULT 1
A59000
mater protein [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: A59000
R;Tong, Z.B.; Nelson, L.M.
Endocrinology 140, 3720-3726, 1999
A;Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature o
A;Reference number: A59000; MUID:99360614; PMID:10433232
A;Accession: A59000
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1111 <KUR>
A;Cross-references: UNIPROT:Q9R1M5; UNIPARC:UPI0000021410; GB:AF074018; NID:g5802697; P
C;Genetics:
A;Gene: Mater

Query Match 48.9%; Score 981; DB 2; Length 1111;
Best Local Similarity 54.5%; Pred. No. 3.1e-77;
Matches 210; Conservative 55; Mismatches 114; Indels 6; Gaps 5;

```
Qy          4 VTEIMSRPERLLFIIDGFDDLGSVL-NNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFL 62
              ||:|||:||||||:||| ||: ||| ::|  | :|| ::|| : |: ||||| |||:|||
Db        257 VTKIMSQPERLLFVIDGLDDMDSVLQHDDMTLSRDWKDEQPIYILMYSLLRKALLPQSFL 316

Qy         63 IVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRELL 122
              |:| |: | ||||| |||| |:|| |:|  :|  |:||     |  : |   ::: | :|
Db        317 IITTRNTGLEKLKSMVVSPLYILVEGLSASRRSQLVLENISNESDRIQVFHSLIENHQLF 376

Qy        123 DQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLEER 182
              |||| |:| ||:| ||||| :|:      ||||||:|  |||||| :   :  |: ||:
Db        377 DQCQAPSVCSLVCEALQLQKKLGKRCTLPCQTLTGLYATLVFHQLTLKRPSQSALSQEEQ 436

Qy        183 VVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSH-CEEYYTFFH 241
              : |   | || ||||  :||| |||     | |||: |||||||||   |  |: | | |
Db        437 ITLVGLCMMAAEGVWTMRSVFYDDDLKNYSLKESEILALFHMNILLQVGHNSEQCYVFSH 496

Qy        242 LSLQDFCAALYYVLEGL-EIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300
              |||||| |||||||||| |      |  ::|  :  ||:|:       | ||||||||:::
Db        497 LSLQDFFAALYYVLEGLEEWNQHFC--FIENQRSIMEVKRTD-DTRLLGMKRFLFGLMNK 553

Qy        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
              |: : ||||   ||   |:||| |||||: || | |:| ||||||:||||:||:|||  |
Db        554 DILKTLEVLFEYPVIPTVEQKLQHWVSLIAQQVNGTSPMDTLDAFYCLFESQDEEFVGGA 613

Qy        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
              |  |||||| ||| :||  ||:||:
Db        614 LKRFQEVWLLINQKMDLKVSSYCLK 638
```

RESULT 2

```
T17255
hypothetical protein DKFZp586O1822.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17255
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17255
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1192 <KOE>
A;Cross-references: UNIPROT:Q9C000; UNIPARC:UPI000016AC73; EMBL:AL117470
A;Experimental source: adult uterus; clone DKFZp586O1822
C;Genetics:
A;Note: DKFZp586O1822.1

  Query Match             24.2%;  Score 484.5;  DB 2;  Length 1192;
  Best Local Similarity   31.9%;  Pred. No. 8.3e-34;
  Matches  127;  Conservative   73;  Mismatches  153;  Indels   45;  Gaps   11;

Qy          2 APVTEIMSRPERLLFIIDGFDDLGSVLNN-DTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              ||: :|:|||||||||:|| |: | ||    ::||  |:: ||   |: ||| | :|||:
Db        112 APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILPEA 171

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
                ::| |    : |   :   |:: | | |   |          | |  :  | : :|:|
Db        172 SFLITARTTALQNLIPSLEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKE 231

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTG--LHAAFVFHQLTPRGVVRRCLN 178
              |   | ||| | | | | |    | :  ::| |   ||       |  | |
Db        232 LWALCLVPWVSWLACTCLMQQMKRKEKLTLTSKTTTTLCLHYLAQALQAQPLG------- 284

Qy        179 LEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNIL----LPDSHCE 234
                    |:  | :| ||:| :|::|  |||   ||  : :     | ||    :| |
Db        285 ----PQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPIPLS--- 337

Qy        235 EYYTPFHLSLQDFCAALYYVLEGLE--IEPALCPLYVEKTKRSMELKQAGFHIHSLW--- 289
                |:| ||  |:| ||: |||| :   : : | : :|||  :       : || |:
Db        338 --YSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEA-------YGIHGLFGAS 388

Qy        290 MKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLG--QQPNATTPGDTLDAFHC 347
              ||| ||:|::  | :| :  | :  |  : |: ||  |    ||:      :|:: ||
Db        389 TTRFLLGLLSDEGEREMENIFHCRLSQG--RNLMQWVPSLQLLLQPH------SLESLHC 440

Qy        348 LFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQ 385
              |:||::| |:   :  |:|: : :  :::|:  :||::
Db        441 LYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIK 478


RESULT 3
S27880
Nasopressin receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
C;Accession: S27880
R;Herrera, V.L.; Ruiz-Opazo, N.
submitted to the EMBL Data Library, February 1992
A;Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
A;Reference number: S27880
A;Accession: S27880
A;Molecule type: mRNA
A;Residues: 1-483 <HER>
A;Cross-references: UNIPARC:UPI0000167960; EMBL:M85183; NID:g202805; PIDN:AAA03623.1; PI
C;Genetics:
A;Gene: AVP

  Query Match             11.1%;  Score 223.5;  DB 2;  Length 483;
  Best Local Similarity   29.6%;  Pred. No. 1.6e-11;
  Matches   82;  Conservative   37;  Mismatches  109;  Indels   49;  Gaps   10;

Qy        144 VGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVF 203
              :|  ::  ::| | ::  |:   |   |     :  |    |:  ||:| ||:   :: |
Db          3 LGRDLSRTSKTTTSVYLLFITSMLKSAGTNGPRVQGE----LRMLCRLAREGILKHQAQF 58

Qy        204 DGDD-----LMVQGLGESELRALFHMNILLPD-SHCEEYYTFFHLSLQDFCAALYYVLEG 257
                  |     | :||   |::: :|      ||       | |   | |:| ||| |:|:
Db         59 SEKDLERLKLKLQG---SQVQTMFLSKKELPGVLETVVTYQFIDQSFQEFLAALSYLLDA 115

Qy        258 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 317
              | |      |:    |     ||   |     ||||||:| :  | :    || ||
Db        116 -EGAPGNSAGSVQMLLNS----DAGLRGHLALTTRFLFGLLSTERIRDIGNHFGCVVPGR 170

Qy        318 VKQKLLHWVSLLGQ-QPNATTPGDT-------------------------LDAFHCLFE 350
              ||| | ||   || ||   | |                           |:  :||:|
Db        171 VKQDTLRWVQ--GQSQPKVATVGAEKKDELKDEEAEEEEEEEEEEEELNFGLELLYCLYE 228

Qy        351 TQDKEFVRLALNSFQEVWLPINQ--NLDLIASSFCLQ 385
              ||: :||| ||:|  |: |   :   :||   |:|:|
Db        229 TQEDDFVRQALSSLPEMVLERVRLTRMDLEVLSYCVQ 265


RESULT 4
A48843
MHC class II transactivator - human
N;Alternate names: CIITA
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: A48843
R;Steimle, V.; Otten, L.A.; Zufferey, M.; Mach, B.
Cell 75, 135-146, 1993
A;Title: Complementation cloning of an MHC class II transactivator mutated in hereditar
A;Reference number: A48843; MUID:94006536; PMID:8402893
A;Accession: A48843
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1130 <STE>
A;Cross-references: UNIPROT:P33076; UNIPARC:UPI00000314FD; GB:X74301; NID:g414112; PIDN

  Query Match             8.5%;  Score 171;  DB 2;  Length 1130;
  Best Local Similarity   23.6%;  Pred. No. 1.9e-06;
  Matches   96;  Conservative   48;  Mismatches  166;  Indels   96;  Gaps   16;

Qy          5 TEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFT----------------LI 48
              : |: ||:|:| |:| |::|                | |  |               |:
Db        485 SHILKRPDRVLLILDAFEEL--------------EAQDGFLHSTCGPAPAEPCSLRGLL 529

Qy         49 RSLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQR----IHLLLERGIG 104
                | :| ||    |::| |  |   ::|   :      : | | ||     :      |:
Db        530 AGLFQKKLLRGCTLLLTARPRG-RLVQSLSKADALFELSGFSMEQAQAYVMRYFESSGMT 588

Qy        105 EHQKTQGLRAIMNNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVF 164
              |||  : |  : :   ||     | :   :|   :   :||  |     |||||:
Db        589 EHQ-DRALTLLRDRPLLLSHSHSPTLCRAVCQLSEALLELGED-AKLPSTLTGLYV---- 642

Qy        165 HQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDL---------MVQGLGE 215
                    |:: |          |     ::| |     :|    |              | :|| :
Db        643 ------GLLGRAALDSPPGALAELAKLAWELGRRHQSTLQEDQFPSADVRTWAMAKGLVQ 696

Qy        216 SELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRS 275
                ||          |     |  || |  ||:  | | ||:     | |:  | |
Db        697 HPPRA-----------AESELAFPSFLLQCFLGALWLALSG-EIKDKELPQYLALTPRK 743

Qy        276 MELKQAGFHIHSLWMK---RFLFGLVSEDVRRPLEVLLGCPVPLGV--KQKLL-HWVSLL 329
                        : |::   ||| ||: :   | |  |||       |  |||:|  ::  |
Db        744 K-------RPYDNWLEGVPRFLAGLIFQPPARCLGALLGPSAAASVDRKQKVLARYLKRL 796

Qy        330 GQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNL 375
                ||        |:  ||  | ::        :|  :  | |
Db        797 --QPGTLRARQLLELLHCAHEAEE----------AGIWQHVVQEL 829


RESULT 5
T31668
hypothetical protein COS1.5 - sea squirt (Ciona intestinalis)
C;Species: Ciona intestinalis
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:21:18 ; Search time 69.981 Seconds
(without alignments)
2417.239 Million cell updates/sec

Title: US-10-677-943A-2
Perfect score: 2005
Sequence: 1 QAPVTEIMSRPERLLFIIDG.....EVWLPINQNLDLIASSFCLQ 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2005	100.0	385	5	AAU79523	Aau79523 Human MAT
2	2005	100.0	476	8	ADP47894	Adp47894 Human CAT
3	2005	100.0	682	8	ADJ19405	Adj19405 Human PAN
4	2005	100.0	1033	5	ABU65214	Abu65214 Human NOV
5	2005	100.0	1033	8	ADN62079	Adn62079 Human nov
6	2005	100.0	1143	6	ADA45221	Ada45221 Human MAT
7	2005	100.0	1162	5	AAE25053	Aae25053 Human EMB
8	2005	100.0	1162	6	ADA45219	Ada45219 Human MAT
9	2005	100.0	1344	5	AAO15585	Aao15585 Human PYR
10	2005	100.0	1344	7	ADF94745	Adf94745 Human PYR
11	2005	100.0	2312	5	AAO17859	Aao17859 Pyrin dom
12	1997	99.6	1200	5	AAU79526	Aau79526 Human MAT
13	1997	99.6	1200	6	AAE31749	Aae31749 Human MAT
14	1997	99.6	1200	9	ADV04681	Adv04681 NALP5, SE
15	1940	96.8	385	6	AAE31746	Aae31746 Human MAT
16	981	48.9	1111	5	AAU79525	Aau79525 Mouse MAT
17	931	46.4	1111	6	AAE31748	Aae31748 Mouse MAT
18	551.5	27.5	612	8	ADJ19397	Adj19397 Murine PA
19	544	27.1	924	8	ADO50890	Ado50890 Human NOV
20	543	27.1	474	8	ADP47888	Adp47888 Human CAT
21	543	27.1	628	8	ADJ19407	Adj19407 Human PAN
22	543	27.1	732	5	AAO17858	Aao17858 Pyrin dom
23	543	27.1	732	8	ADJ19387	Adj19387 Human PAN
24	543	27.1	920	5	ABP53254	Abp53254 Human MDD
25	543	27.1	924	7	ADF44705	Adf44705 Human NOV
26	543	27.1	1012	8	ADP47734	Adp47734 Human inf
27	543	27.1	1093	9	ADV44750	Adv44750 Human nuc
28	541.5	27.0	674	8	ADJ19398	Adj19398 Human NAL
29	541.5	27.0	717	8	ADP47754	Adp47754 Human CAT
30	541.5	27.0	719	5	ADE36452	Ade36452 Human PAA
31	541.5	27.0	719	8	ADJ19374	Adj19374 Human PAN
32	541.5	27.0	920	7	ADM34360	Adm34360 Human mut
33	541.5	27.0	944	8	ADP47756	Adp47756 Human CAT
34	541.5	27.0	977	8	ADP47870	Adp47870 Human CAT
35	541.5	27.0	1034	4	AAE07514	Aae07514 Human PYR
36	541.5	27.0	1034	6	ABU08503	Abu08503 Huma PYRI
37	541.5	27.0	1034	6	ABU63315	Abu63315 Human pyr
38	541.5	27.0	1034	7	ADM34355	Adm34355 Human mut
39	541.5	27.0	1034	7	ADM34351	Adm34351 Human mut
40	541.5	27.0	1034	7	ADM34319	Adm34319 Human cry
41	541.5	27.0	1034	8	ADP24014	Adp24014 PRO polyp
42	539.5	26.9	472	8	ADP47889	Adp47889 Human CAT
43	538.5	26.9	479	8	ADP47883	Adp47883 Human CAT
44	538.5	26.9	1034	7	ADM34353	Adm34353 Human mut
45	538.5	26.9	1034	7	ADM34354	Adm34354 Human mut

ALIGNMENTS

RESULT 1
AAU79523
ID AAU79523 standard; protein; 385 AA.
XX
AC AAU79523;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER peptide encoded by cDNA fragment #1.
XX
KW Human; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; gene therapy; fertility.
XX
OS Homo sapiens.
XX
PN WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI; 2002-454595/48.
DR N-PSDB; ABK48609.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 1; Page 69-70; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,

CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The protein sequence presented is human MATER
CC which is encoded by cDNA fragment #1, isolated from a human cDNA library
XX
SQ Sequence 385 AA;

Query Match 100.0%; Score 2005; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.9e-209;
Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240

Qy        241 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300

Qy        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360

Qy        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
              |||||||||||||||||||||||||
Db        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
```

RESULT 2
ADP47894
ID ADP47894 standard; protein; 476 AA.
XX
AC ADP47894;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human CATERPILLER 19.8 nucleotide binding domain (NBD) protein.
XX
KW Monarch-1; CATERPILLER 11.2; caspase recruitment domain;
KW CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat;
KW CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIAS1;
KW cold-induced autoinflammatory syndrome 1; antiinflammatory; cytostatic;
KW inflammatory disease; cancer; gene therapy; human; purine;
KW nucleotide binding domain; NBD; CATERPILLER 19.8.
XX
OS Homo sapiens.
XX
PN WO2004034093-A2.
XX
PD 22-APR-2004.
XX
PF 30-APR-2003; 2003WO-US013562.
XX
PR 30-APR-2002; 2002US-0376626P.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Ting JY, Linhoff MW, Harton JA, Williams KL, Lich J, O'connor W;
PI Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X;
XX
DR WPI; 2004-348215/32.
XX
PT New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3,
PT CATERPILLAR 16.1, CATERPILLER 16.2 or CIAS1 polypeptide, useful in
PT preparing a composition for treating inflammatory disease or cancer.
XX
PS Example 1; SEQ ID NO 173; 205pp; English.
XX
CC The invention relates to a novel isolated nucleic acid encoding a Monarch
CC -1, CATERPILLER (CARD [caspase recruitment domain], transcription
CC enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2,
CC CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 (cold-
CC induced autoinflammatory syndrome 1) polypeptide comprising the amino
CC acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or
CC its fragment. The nucleic acid of the invention demonstrates
CC antiinflammatory and cytostatic activities and may be useful in preparing
CC a composition for treating an inflammatory disease or cancer, possibly
CC via gene therapy. The current sequence is that of a human CATERPILLER
CC nucleotide binding domain (NBD) protein of the invention.
XX
SQ Sequence 476 AA;

Query Match 100.0%; Score 2005; DB 8; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.6e-209;
Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 QAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPES 139

Qy         61 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 FLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIMNNRE 199

Qy        121 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 LLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFVFHQLTPRGVVRRCLNLE 259

Qy        181 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 ERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFF 319

Qy        241 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 HLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSE 379

Qy        301 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 DVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 439

Qy        361 LNSFQEVWLPINQNLDLIASSFCLQ 385
              |||||||||||||||||||||||||
Db        440 LNSFQEVWLPINQNLDLIASSFCLQ 464
```

RESULT 3
ADJ19405
ID ADJ19405 standard; protein; 682 AA.
XX
AC ADJ19405;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human PAN4 PAAD/NACHT/ARED domain-containing protein - SEQ ID 96.
XX
KW PAAD; cytostatic; cancer; gene therapy; human; PAAD domain; NACHT;
KW NB-ARC; domain; nucleotide binding; ANGIO-R; ARED domain;
KW angiotensin receptor-like; PAN4.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:21:18 ; Search time 218.123 Seconds
(without alignments)
2417.239 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6298	100.0	1200	5	AAU79526	Aau79526 Human MAT
2	6298	100.0	1200	6	AAE31749	Aae31749 Human MAT
3	6298	100.0	1200	9	ADV04681	Adv04681 NALP5, SE
4	6077.5	96.5	2312	5	AAO17859	Aao17859 Pyrin dom
5	5892.5	93.6	1162	5	AAE25053	Aae25053 Human EMB
6	5892.5	93.6	1162	6	ADA45219	Ada45219 Human MAT
7	5804	92.2	1143	6	ADA45221	Ada45221 Human MAT
8	5741.5	91.2	1344	5	AAO15585	Aao15585 Human PYR
9	5741.5	91.2	1344	7	ADF94745	Adf94745 Human PYR
10	5134	81.5	1033	5	ABU65214	Abu65214 Human NOV
11	5125	81.4	1033	8	ADN62079	Adn62079 Human nov
12	3553	56.4	682	8	ADJ19405	Adj19405 Human PAN
13	2747	43.6	1111	5	AAU79525	Aau79525 Mouse MAT
14	2610	41.4	1111	6	AAE31748	Aae31748 Mouse MAT
15	2484	39.4	476	8	ADP47894	Adp47894 Human CAT
16	1997	31.7	385	5	AAU79523	Aau79523 Human MAT
17	1932	30.7	385	6	AAE31746	Aae31746 Human MAT
18	1874.5	29.8	1093	9	ADV44750	Adv44750 Human nuc
19	1659.5	26.3	1012	8	ADP47734	Adp47734 Human inf
20	1516.5	24.1	935	8	ADP47736	Adp47736 Human inf
21	1409.5	22.4	994	5	AAO15593	Aao15593 Human PYR
22	1409.5	22.4	994	5	ADE36409	Ade36409 Human PAA
23	1409.5	22.4	994	7	ADF94761	Adf94761 Human PYR
24	1409.5	22.4	994	8	ADJ19331	Adj19331 Human PAN
25	1408.5	22.4	994	8	ABM82183	Abm82183 Tumour-as
26	1397	22.2	924	7	ADF44705	Adf44705 Human NOV
27	1387	22.0	924	8	ADO50890	Ado50890 Human NOV
28	1379	21.9	258	5	AAU79524	Aau79524 Human MAT
29	1369	21.7	983	5	AAO17870	Aao17870 Pyrin dom
30	1322.5	21.0	1062	8	ADP47724	Adp47724 Human Mon
31	1320	21.0	991	9	ADV44751	Adv44751 Human nuc
32	1320	21.0	1035	5	ADE36417	Ade36417 Human PAA
33	1320	21.0	1035	8	ADJ19339	Adj19339 Human PAN
34	1320	21.0	1099	5	AAO17857	Aao17857 Pyrin dom
35	1319	20.9	1027	8	ADJ19383	Adj19383 Human PAN
36	1319	20.9	1061	5	AAO15590	Aao15590 Human PYR
37	1319	20.9	1061	7	ADF94755	Adf94755 Human PYR
38	1312	20.8	1036	9	ADV44755	Adv44755 Human nuc
39	1304.5	20.7	919	6	ADA55150	Ada55150 Human pro
40	1303.5	20.7	919	5	ABP43483	Abp43483 Human sec
41	1303.5	20.7	919	6	ABJ19319	Abj19319 NOVX rela
42	1303.5	20.7	919	8	ADO41638	Ado41638 Novel hum
43	1301	20.7	258	6	AAE31747	Aae31747 Human MAT
44	1278.5	20.3	858	5	AAM50328	Aam50328 Human nuc
45	1252.5	19.9	1006	8	ADP47726	Adp47726 Human Mon

ALIGNMENTS

RESULT 1
AAU79526
ID AAU79526 standard; protein; 1200 AA.
XX
AC AAU79526;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER protein.
XX
KW Human; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
XX
PN WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI; 2002-454595/48.
DR N-PSDB; ABK48628.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 1; Fig 9; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,

```
CC   the MATER protein is useful as a contraceptive agent, or for influencing
CC   (either inhibiting or enhancing) fertility and can be used to detect a
CC   predisposition to infertility or reduced fertility, or for presymptomatic
CC   screening of an individual for infertility/reduced fertility. The protein
CC   and polynucleotide are also useful for detecting an excess or deficiency,
CC   or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC   human or a mouse) or for screening for a compound useful in influencing
CC   MATER-mediated fertility. The protein sequence presented is the human
CC   MATER. Note: Residues 1-25 are encoded by nucleotides which were deduced,
CC   by the inventors, by comparison to published human genomic DNA sequences,
CC   while the remainder of the sequence was determined by direct cloning of
CC   human ovarian cDNAs
XX
SQ   Sequence 1200 AA;

  Query Match             100.0%;  Score 6298;  DB 5;  Length 1200;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780

Qy        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840

Qy        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900

Qy        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960

Qy        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020

Qy       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080

Qy       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140

Qy       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
```

```
RESULT 2
AAE31749
ID   AAE31749 standard; protein; 1200 AA.
XX
AC   AAE31749;
XX
DT   07-MAR-2003  (first entry)
XX
DE   Human MATER protein.
XX
KW   Human; MATER protein; infertility; fertility; contraceptive agent;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200281492-A1.
XX
PD   17-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009776.
XX
PR   04-APR-2001; 2001WO-US010981.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nelson LM,  Tong Z;
XX
DR   WPI; 2003-058494/05.
DR   N-PSDB; AAD49018.
XX
PT   New isolated variant MATER proteins and nucleic acids, useful for
PT   diagnosing, prognosing and treating infertility and reduced fertility,
PT   and as contraceptive agents.
XX
PS   Claim 1; Page 99-104; 110pp; English.
XX
CC   The present invention relates to novel MATER proteins and polynucleotides
CC   encoding such proteins. The MATER proteins are essential to fertility.
CC   Sequences of the invention are useful for diagnosing, prognosing and
CC   treating infertility, reduced fertility and as contraceptive agents. They
CC   are also useful in gene therapy. The method is useful for detecting a
CC   predisposition to or pre-symptomatic screening of an individual for
CC   infertility or reduced fertility. The present sequence is human MATER
CC   protein
XX
SQ   Sequence 1200 AA;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:22:26 ; Search time 46.229 Seconds
(without alignments)
2497.570 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

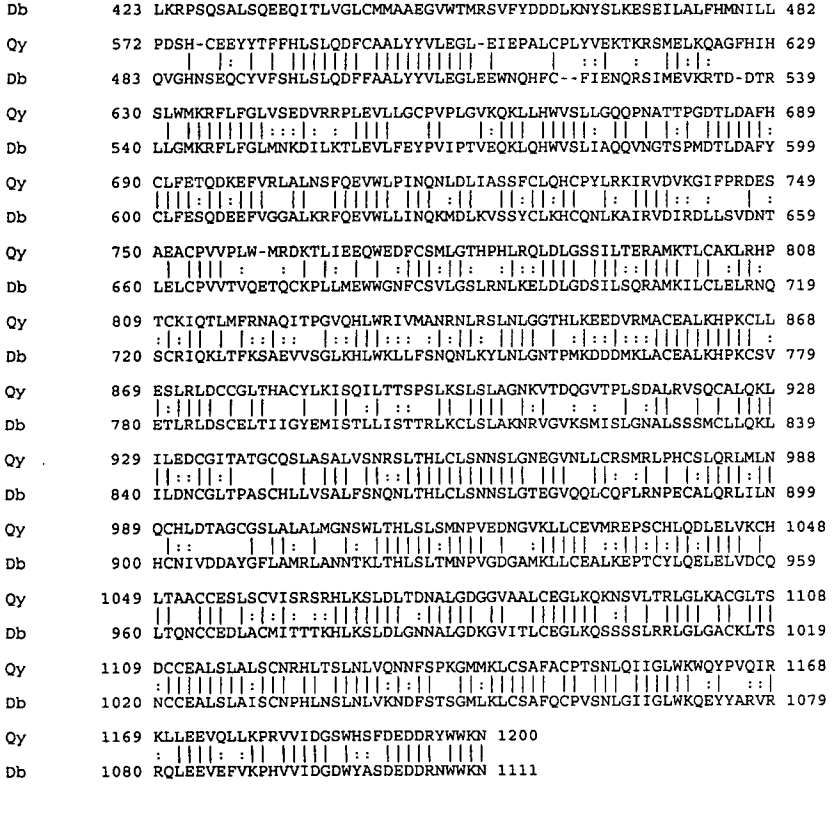
Result No.	Score	% Query Match	Length	DB	ID	Description
1	2747	43.6	1111	2	A59000	mater protein [imp
2	914	14.5	1192	2	T17255	hypothetical prote
3	633.5	10.1	456	2	S20597	ribonuclease inhib
4	624	9.9	456	2	A31857	ribonuclease inhib
5	611.5	9.7	461	2	A31858	ribonuclease-angio
6	367.5	5.8	483	2	S27880	Nasopressin recept
7	327	5.2	1130	2	A48843	MHC class II trans
8	220	3.5	631	2	C89243	protein F28C1.3 [i
9	220	3.5	631	2	T21471	hypothetical prote
10	205	3.3	1075	2	T31668	hypothetical prote
11	201.5	3.2	545	2	T52068	RAN GTPase-activat
12	197	3.1	1004	2	T31665	hypothetical prote
13	192.5	3.1	568	2	F86291	hypothetical prote
14	188.5	3.0	1167	2	T06146	disease resistance
15	185	2.9	533	2	T52063	ran GTPase-activat
16	183	2.9	535	2	T48102	RAN GTPase activat
17	179.5	2.9	607	2	E96598	protein F20N2.2 [i
18	177	2.8	789	2	T52067	hypothetical prote
19	175.5	2.8	506	2	A45841	T-complex-associat
20	172	2.7	907	2	F96617	probable disease r
21	172	2.7	1039	2	A85096	hypothetical prote
22	171	2.7	934	2	T05201	hypothetical prote
23	168.5	2.7	1184	2	H71436	hypothetical prote
24	168.5	2.7	1301	2	D85188	disease resistance
25	167	2.7	712	2	C71419	hypothetical prote
26	164.5	2.6	1063	2	E96662	hypothetical prote
27	163.5	2.6	1317	2	B85189	disease resistance
28	162.5	2.6	998	2	T23427	hypothetical prote
29	161	2.6	1422	2	B71437	probable resistanc
30	159.5	2.5	1253	2	T45787	disease resistance
31	158.5	2.5	1240	2	T06404	resistance complex
32	157.5	2.5	1152	2	T31911	hypothetical prote
33	155.5	2.5	1219	2	T06608	disease resistance
34	153.5	2.4	618	2	T48193	hypothetical prote
35	151.5	2.4	1239	2	T06143	disease resistence
36	150.5	2.4	630	2	S13724	ESAG 8 protein - T
37	150.5	2.4	1017	2	H96663	hypothetical prote
38	150	2.4	271	2	D84586	hypothetical prote
39	148.5	2.4	587	2	JC5300	Ran GTPase activat
40	148	2.3	1304	2	G85188	disease resistance
41	147.5	2.3	966	2	D96662	hypothetical prote
42	147	2.3	589	2	T52070	RNA1 protein homol
43	146	2.3	589	2	A36983	RNA1 homolog fug1
44	145	2.3	1121	2	T02764	myosin-I binding p
45	145	2.3	1189	2	T52346	disease resistance

ALIGNMENTS

RESULT 1
A59000
mater protein [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: A59000
R;Tong, Z.B.; Nelson, L.M.
Endocrinology 140, 3720-3726, 1999
A;Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature o
A;Reference number: A59000; MUID:99360614; PMID:10433232
A;Accession: A59000
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1111 <KUR>
A;Cross-references: UNIPROT:Q9R1M5; UNIPARC:UPI0000021410; GB:AF074018; NID:g5802697; P
C;Genetics:
A;Gene: Mater

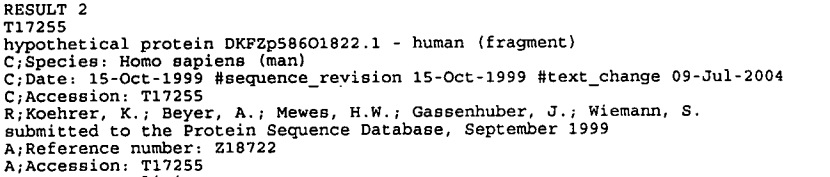
Query Match 43.6%; Score 2747; DB 2; Length 1111;
Best Local Similarity 50.7%; Pred. No. 2.6e-176;
Matches 564; Conservative 185; Mismatches 307; Indels 56; Gaps 14;

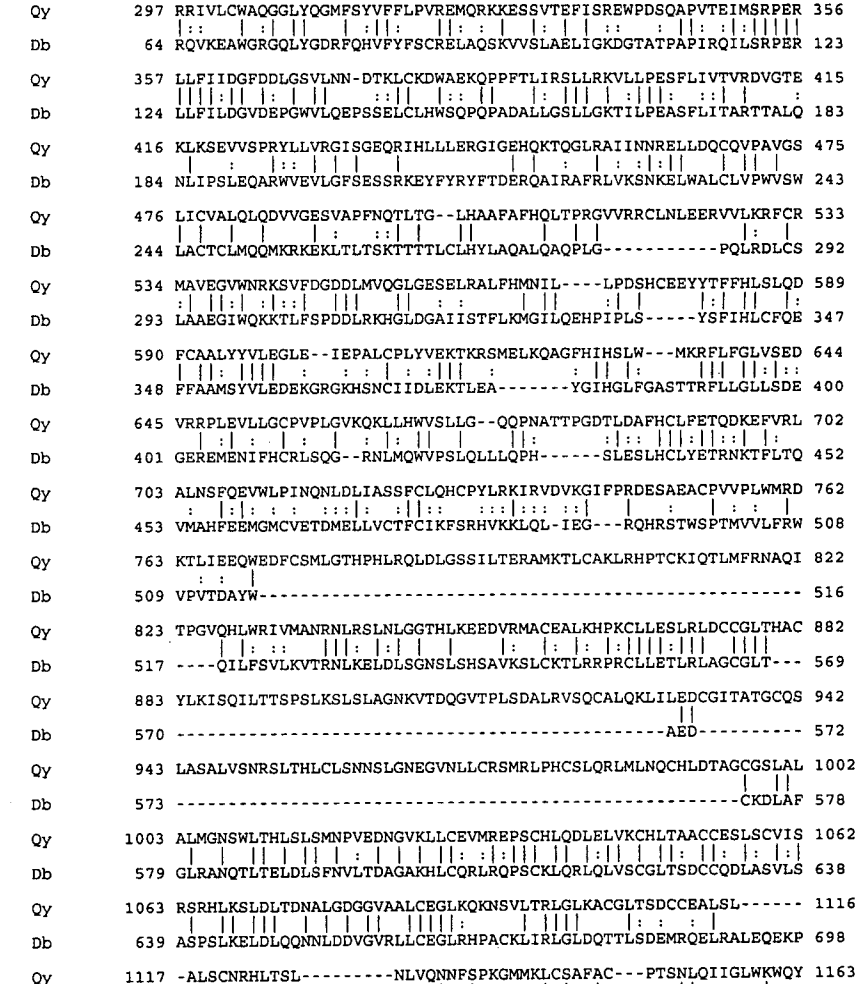
```
Qy      133 MNLRTLSEKARDDMKRHSPE-DPEATMTDQGPSKE-----------------KVPGISQ 173
            :  | | |: : : |  ||| | ::   |||| :|                   |  |: :
Db       12 LKARGLEEEQKSERKMTSPENDSKSIQKDQGPEQEQTSESTMGPPEKDSKAILKARGLEE 71

Qy      174 AVQQDS----------ATAAETKEQEISQAMEQEGATAAETEEQEISQ----AMEQEGAT 219
              : :|           |   ::  :|: || |::  |: | : : | :      ||   |
Db       72 EQKSESTMSPSENVSRAILKDSGSEEVEQASERK-MTSPENDSKSIQKDQGPEQEQTSET 130

Qy      220 AAETEE-------QGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDS 272
              ||        : :|||  |||:||: ||     |:         | |||: |: ||
Db      131 LQSKEEDEVTEADKDNGGDLQDYKAHVIAKFDTSVDLH-------YDSPEMKLLSDAFKP 183

Qy      273 DRWGFRPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSV 332
            :  |:| |::|||: |:|||||| ||| |||| |:| | |:| |  |||::  ::||:
Db      184 YQKTFQPHTIILHGRPGVGKSALARSIVLGWAQGKLFQKM-SFVIFFSVREIKWTEKSSL 242

Qy      333 TEFISREWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVL-NNDTKLCKDWAEKQPPFTLI 391
            : |::| |||   ||:|||:|||||:||| ||: ||| ::|  | :|| ::|| : |:
Db      243 AQLIAKECPDSWDLVTKIMSQPERLLFVIDGLDDMDSVLQHDDMTLSRDWKDEQPIYILM 302

Qy      392 RSLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQK 451
            ||||| |||:||||:| |: | ||||| |||| |:|| |:|  :|  :|  |:||    |  :
Db      303 YSLLRKALLPQSFLIITTRNTGLEKLKSMVVSPLYILVEGLSASRRSQLVLENISNESDR 362

Qy      452 TQGLRAIINNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLT 511
            |   ::| | :| |||| |:| ||:| ||||  :|:      ||||||:|   |||||
Db      363 IQVFHSLIENHQLFDQCQAPSVCSLVCEALQLQKKLGKRCTLPCQTLTGLYATLVFHQLT 422

Qy      512 PRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILL 571
              :  : |: ||:: |   | || ||||  :|||  |||    | |||: |||||||||
```



RESULT 2
T17255
hypothetical protein DKFZp586O1822.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17255
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17255
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1192 <KOE>
A;Cross-references: UNIPROT:Q9C000; UNIPARC:UPI000016AC73; EMBL:AL117470
A;Experimental source: adult uterus; clone DKFZp586O1822
C;Genetics:
A;Note: DKFZp586O1822.1

Query Match 14.5%; Score 914; DB 2; Length 1192;
Best Local Similarity 27.3%; Pred. No. 5.1e-53;
Matches 274; Conservative 142; Mismatches 314; Indels 274; Gaps 28;





RESULT 3
S20597
ribonuclease inhibitor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S20597
R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:38:41 ; Search time 183.614 Seconds
(without alignments)
2730.707 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6298	100.0	1200	4	US-10-399-443-24	Sequence 24, Appl
2	6298	100.0	1200	4	US-10-677-943-24	Sequence 24, Appl
3	6298	100.0	1200	5	US-10-860-761-4	Sequence 4, Appli
4	5892.5	93.6	1162	4	US-10-216-645-2	Sequence 2, Appli
5	5892.5	93.6	1162	4	US-10-416-642-1	Sequence 1, Appli
6	5804	92.2	1143	4	US-10-216-645-4	Sequence 4, Appli
7	5741.5	91.2	1344	4	US-10-124-498-6	Sequence 6, Appli
8	5741.5	91.2	1344	4	US-10-066-521-6	Sequence 6, Appli
9	5134	81.5	1033	4	US-10-092-900A-348	Sequence 348, App
10	3553	56.4	682	4	US-10-407-866-96	Sequence 96, Appl
11	2747	43.6	1111	4	US-10-216-645-5	Sequence 5, Appli
12	2610	41.4	1111	4	US-10-399-443-6	Sequence 6, Appli
13	2610	41.4	1111	4	US-10-677-943-6	Sequence 6, Appli
14	1932	30.7	385	4	US-10-399-443-2	Sequence 2, Appli
15	1932	30.7	385	4	US-10-677-943-2	Sequence 2, Appli
16	1874.5	29.8	1093	5	US-10-794-342-13	Sequence 13, Appl
17	1409.5	22.4	994	3	US-09-965-621-16	Sequence 16, Appl
18	1409.5	22.4	994	4	US-10-124-498-24	Sequence 24, Appl
19	1409.5	22.4	994	4	US-10-066-521-24	Sequence 24, Appl
20	1409.5	22.4	994	4	US-10-407-866-16	Sequence 16, Appl
21	1409.5	22.4	994	4	US-10-781-294-16	Sequence 16, Appl
22	1397	22.2	924	4	US-10-357-820-52	Sequence 52, Appl
23	1320	21.0	991	5	US-10-794-342-14	Sequence 14, Appl
24	1320	21.0	1035	3	US-09-965-621-24	Sequence 24, Appl
25	1320	21.0	1035	4	US-10-407-866-24	Sequence 24, Appl
26	1320	21.0	1035	4	US-10-781-294-24	Sequence 24, Appl
27	1319	20.9	1027	4	US-10-407-866-68	Sequence 68, Appl
28	1319	20.9	1061	4	US-10-124-498-18	Sequence 18, Appl
29	1319	20.9	1061	4	US-10-066-521-18	Sequence 18, Appl
30	1312	20.8	1036	5	US-10-794-342-18	Sequence 18, Appl
31	1304.5	20.7	919	4	US-10-094-749-2718	Sequence 2718, Ap
32	1303.5	20.7	919	4	US-10-161-493-4	Sequence 4, Appli
33	1303.5	20.7	919	4	US-10-398-037-7	Sequence 7, Appli
34	1301	20.7	258	4	US-10-399-443-4	Sequence 4, Appli
35	1301	20.7	258	4	US-10-677-943-4	Sequence 4, Appli
36	1278.5	20.3	858	3	US-09-848-035-8	Sequence 8, Appli
37	1278.5	20.3	858	3	US-09-986-224-8	Sequence 8, Appli
38	1278.5	20.3	858	5	US-10-116-422-8	Sequence 8, Appli
39	1252.5	19.9	1032	5	US-10-882-761-33	Sequence 33, Appl
40	1252.5	19.9	1034	4	US-10-127-516-5	Sequence 5, Appli
41	1252.5	19.9	1034	4	US-10-027-629-5	Sequence 5, Appli
42	1252.5	19.9	1034	4	US-10-028-374-18	Sequence 18, Appl
43	1252.5	19.9	1034	4	US-10-132-967-5	Sequence 5, Appli
44	1252.5	19.9	1034	4	US-10-183-770-18	Sequence 18, Appl
45	1252.5	19.9	1034	4	US-10-264-958B-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-10-399-443-24
; Sequence 24, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-443-24

Query Match 100.0%; Score 6298; DB 4; Length 1200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780

Qy        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840

Qy        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900

Qy        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960

Qy        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020

Qy       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080

Qy       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140

Qy       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
```

```
RESULT 2
US-10-677-943-24
; Sequence 24, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
;  APPLICANT: The Government of the United States of America as
;  APPLICANT:  represented by the Secretary of the Department of Health and
;  APPLICANT:  Human Services
;  APPLICANT:  Nelson, Lawrence
;  APPLICANT:  Tong, Zhi-Bin
;  TITLE OF INVENTION: Human Gene Critical to Fertility
;  FILE REFERENCE: 4239-64790
;  CURRENT APPLICATION NUMBER: US/10/677,943
;  CURRENT FILING DATE:  2003-10-01
;  PRIOR APPLICATION NUMBER: 60/241,510
;  PRIOR FILING DATE: 2000-10-18
;  PRIOR APPLICATION NUMBER: PCT/US02/09776
;  PRIOR FILING DATE: 2002-03-29
;  PRIOR APPLICATION NUMBER: PCT/US01/10981
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
;   LENGTH: 1200
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-677-943-24

  Query Match           100.0%;  Score 6298;  DB 4;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:22:06 ; Search time 235.052 Seconds
(without alignments)
3601.908 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6298	100.0	1200	1	NALP5_HUMAN	P59047 homo sapien
2	3106.5	49.3	1098	2	Q647I9_BOVIN	Q647i9 bos taurus
3	2754	43.7	1111	2	Q6XZF2_MOUSE	Q6xzf2 mus musculu
4	2753	43.7	1095	2	Q6VSG3_MOUSE	Q6vsg3 mus musculu
5	2747	43.6	1111	1	NALP5_MOUSE	Q9r1m5 mus musculu
6	2747	43.6	1111	2	Q53ZD5_MOUSE	Q53zd5 mus musculu
7	2747	43.6	1111	2	Q7TPU9_MOUSE	Q7tpu9 mus musculu
8	2746	43.6	1095	2	Q6VSG5_MOUSE	Q6vsg5 mus musculu
9	2745	43.6	1111	2	Q6VSG0_MOUSE	Q6vsg0 mus musculu
10	2626	41.7	1085	2	Q6VSG1_MOUSE	Q6vsg1 mus musculu
11	2262	35.9	989	2	Q6VSG4_MOUSE	Q6vsg4 mus musculu
12	1874.5	29.8	1093	1	NAL14_HUMAN	Q86w24 homo sapien
13	1525	24.2	727	2	Q6VSG2_MOUSE	Q6vsg2 mus musculu
14	1515	24.1	727	2	Q6VSF9_MOUSE	Q6vsf9 mus musculu
15	1417.5	22.5	993	2	Q6B966_MOUSE	Q6b966 mus musculu
16	1409.5	22.4	994	1	NALP4_HUMAN	Q96mn2 homo sapien
17	1408.5	22.4	994	2	Q86W87_HUMAN	Q86w87 homo sapien
18	1404.5	22.3	986	2	Q66X14_MOUSE	Q66x14 mus musculu
19	1355	21.5	996	2	Q4U4S5_BOVIN	Q4u4s5 bos taurus
20	1330	21.1	1043	1	NAL13_HUMAN	Q86w25 homo sapien
21	1320	21.0	991	2	Q7RTR0_HUMAN	Q7rtr0 homo sapien
22	1319	20.9	1061	1	NAL12_HUMAN	P59046 homo sapien
23	1289.5	20.5	986	2	Q86W27_HUMAN	Q86w27 homo sapien
24	1252.5	19.9	1034	1	CIAS1_HUMAN	Q96p20 homo sapien
25	1250	19.8	1033	1	CIAS1_MOUSE	Q8r4b8 mus musculu
26	1245.5	19.8	982	2	Q66X27_MOUSE	Q66x27 mus musculu
27	1242	19.7	1062	2	Q53FL5_HUMAN	Q53fl5 homo sapien
28	1241	19.7	1062	1	NALP2_HUMAN	Q9nx02 homo sapien
29	1226.5	19.5	978	2	Q66X19_MOUSE	Q66x19 mus musculu
30	1178	18.7	982	2	Q8BU40_MOUSE	Q8bu40 mus musculu
31	1145	18.2	1048	2	Q7RTR4_HUMAN	Q7rtr4 homo sapien
32	1128	17.9	1003	2	Q66X22_MOUSE	Q66x22 mus musculu
33	1115	17.7	1009	2	Q7RTR1_HUMAN	Q7rtr1 homo sapien
34	1097	17.4	980	1	NALP7_HUMAN	Q8wx94 homo sapien
35	1082.5	17.2	949	2	Q66X03_MOUSE	Q66x03 mus musculu
36	1082.5	17.2	1029	2	Q86W28_HUMAN	Q86w28 homo sapien
37	1060	16.8	1033	1	NAL11_HUMAN	P59045 homo sapien
38	1060	16.8	1033	2	Q53ZZ0_HUMAN	Q53zz0 homo sapien
39	1047	16.6	1004	2	Q66X01_MOUSE	Q66x01 mus musculu
40	983	15.6	1473	1	NALP1_HUMAN	Q9c000 homo sapien
41	981	15.6	1375	2	Q86UB5_HUMAN	Q86ub5 homo sapien
42	954.5	15.2	959	2	Q66X05_MOUSE	Q66x05 mus musculu
43	951	15.1	934	2	Q66X24_MOUSE	Q66x24 mus musculu
44	944.5	15.0	931	2	Q6UQE7_MOUSE	Q6uqe7 mus musculu
45	940	14.9	863	2	Q8C6J9_MOUSE	Q8c6j9 mus musculu

ALIGNMENTS

RESULT 1
NALP5_HUMAN
ID NALP5_HUMAN STANDARD; PRT; 1200 AA.
AC P59047; Q86W29;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 5 (Mater protein homolog).
GN Name=NALP5; Synonyms=MATER;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RX MEDLINE=21922687; PubMed=11925379; DOI=10.1093/humrep/17.4.903;
RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT "A human homologue of mouse Mater, a maternal effect gene essential
RT for early embryonic development.";
RL Hum. Reprod. 17:903-911(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22451042; PubMed=12563287; DOI=10.1038/nrm1019;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
CC -!- TISSUE SPECIFICITY: Oocyte specific.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 13 LRR (leucine-rich) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC Mater.
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---
DR EMBL; AY054986; AAL15549.1; -; mRNA.
DR EMBL; AY154460; AAO18156.1; -; mRNA.
DR HSSP; P13489; 1A4Y.
DR Ensembl; ENSG00000171487; Homo sapiens.
DR HGNC; HGNC:21269; NALP5.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR Pfam; PF00560; LRR_1; 2.
DR Pfam; PF05729; NACHT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS50824; DAPIN; 1.

```
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Nucleotide-binding; Repeat.
FT   DOMAIN       57    148       DAPIN.
FT   DOMAIN      280    602       NACHT.
FT   REPEAT      704    727       LRR 1.
FT   REPEAT      730    753       LRR 2.
FT   REPEAT      780    803       LRR 3.
FT   REPEAT      809    832       LRR 4.
FT   REPEAT      836    863       LRR 5.
FT   REPEAT      865    892       LRR 6.
FT   REPEAT      893    916       LRR 7.
FT   REPEAT      950    973       LRR 8.
FT   REPEAT      979   1002       LRR 9.
FT   REPEAT     1007   1034       LRR 10.
FT   REPEAT     1036   1059       LRR 11.
FT   REPEAT     1064   1092       LRR 12.
FT   REPEAT     1121   1142       LRR 13.
FT   NP_BIND     286    293       ATP (Potential).
FT   CONFLICT    459    459       I -> M (in Ref. 2).
FT   CONFLICT    506    506       A -> V (in Ref. 2).
FT   CONFLICT    912    912       T -> M (in Ref. 2).
FT   CONFLICT    999    999       S -> F (in Ref. 2).
SQ   SEQUENCE   1200 AA;  134236 MW;  9A070D2A771B28FA CRC64;

  Query Match             100.0%;  Score 6298;  DB 1;  Length 1200;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780

Qy        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840

Qy        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900

Qy        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960

Qy        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020

Qy       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080

Qy       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140

Qy       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
```

```
RESULT 2
Q647I9_BOVIN
ID   Q647I9_BOVIN PRELIMINARY;      PRT;  1098 AA.
AC   Q647I9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Maternal antigen that embryo require.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC   Pecora; Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15189828; DOI=10.1095/biolreprod.104.030288;
RA   Pennetier S., Uzbekova S., Perreau C., Papillier P., Mermillod P.,
RA   Dalbies-Tran R.;
RT   "Spatio-temporal expression of the germ cell marker genes MATER, ZAR1,
RT   GDF9, BMP15,andVASA in adult bovine tissues, oocytes, and
RT   preimplantation embryos.";
RL   Biol. Reprod. 71:1359-1366(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Pennetier S., Uzbekova S., Dalbies-Tran R.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY721594; AAU20764.1; -; mRNA.
DR   InterPro; IPR000887; Aldlse_KDPG_KHG.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN.
DR   Pfam; PF00560; LRR_1; 1.
DR   Pfam; PF05729; NACHT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00368; LRR_RI; 12.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:26:25 ; Search time 57.949 Seconds
(without alignments)
1712.037 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

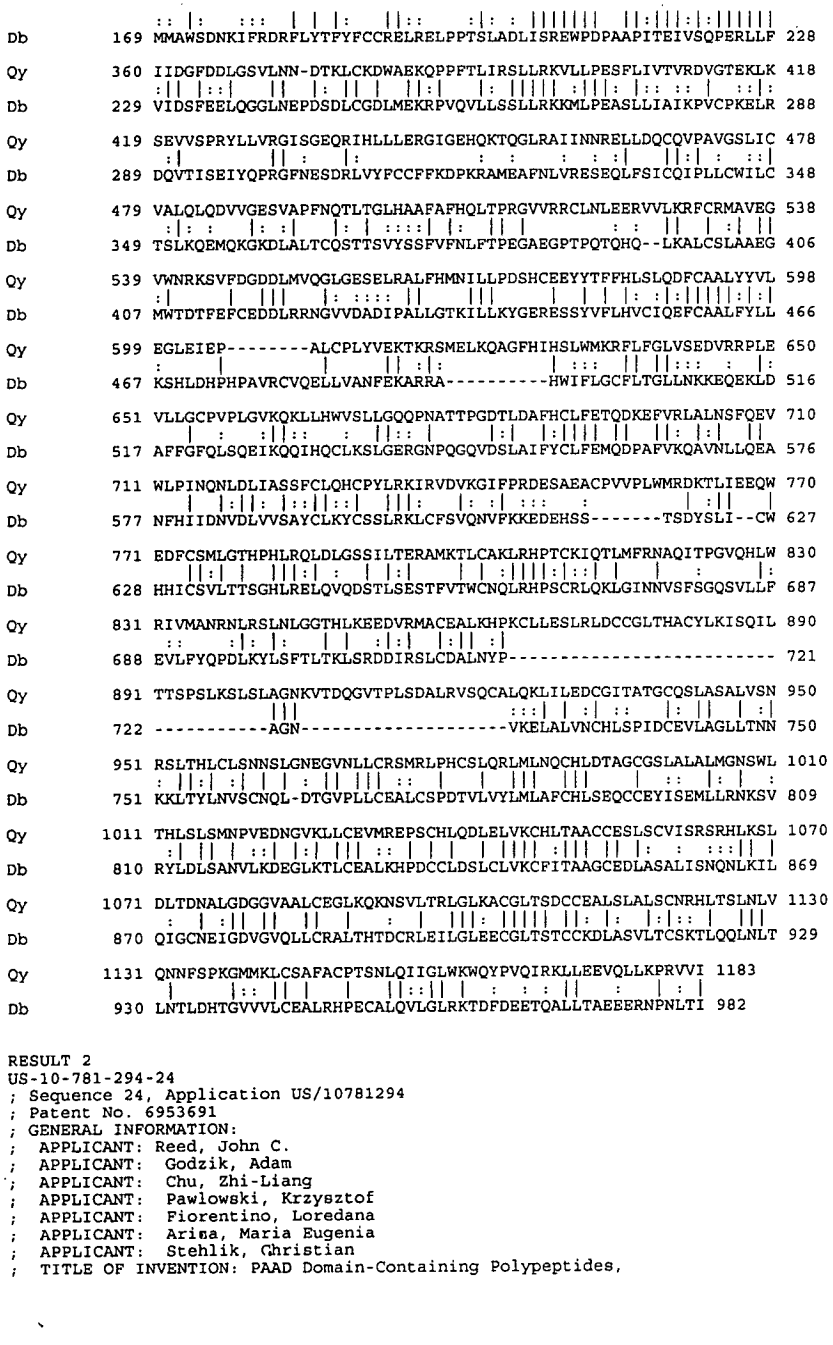
Result No.	Score	% Query Match	Length	DB	ID	Description
1	1409.5	22.4	994	2	US-10-781-294-16	Sequence 16, Appl
2	1320	21.0	1035	2	US-10-781-294-24	Sequence 24, Appl
3	1252.5	19.9	1034	2	US-10-183-770A-18	Sequence 18, Appl
4	1224.5	19.4	1033	2	US-10-029-347-4	Sequence 4, Appli
5	985	15.6	1429	2	US-10-029-347-3	Sequence 3, Appli
6	985	15.6	1429	2	US-10-183-770A-3	Sequence 3, Appli
7	983	15.6	1429	2	US-10-029-347-26	Sequence 26, Appl
8	983	15.6	1429	2	US-10-183-770A-15	Sequence 15, Appl
9	983	15.6	1429	2	US-09-996-617-2	Sequence 2, Appli
10	983	15.6	1473	2	US-09-388-221B-2	Sequence 2, Appli
11	979	15.5	1454	2	US-09-388-221B-10	Sequence 10, Appl
12	976	15.5	1399	2	US-09-388-221B-4	Sequence 4, Appli
13	976	15.5	1443	2	US-09-388-221B-6	Sequence 6, Appli
14	972	15.4	1424	2	US-09-388-221B-12	Sequence 12, Appl
15	876.5	13.9	603	2	US-10-781-294-64	Sequence 64, Appl
16	713.5	11.3	947	2	US-10-781-294-18	Sequence 18, Appl
17	693	11.0	719	2	US-10-781-294-59	Sequence 59, Appl
18	633.5	10.1	456	1	US-08-910-731-4	Sequence 4, Appli
19	633.5	10.1	456	1	US-08-795-395-4	Sequence 4, Appli
20	624	9.9	456	1	US-08-910-731-2	Sequence 2, Appli
21	624	9.9	456	1	US-08-795-395-2	Sequence 2, Appli
22	624	9.9	456	2	US-10-188-495-64	Sequence 64, Appl
23	615	9.8	456	1	US-08-910-731-8	Sequence 8, Appli
24	611.5	9.7	681	2	US-10-781-294-20	Sequence 20, Appl
25	610.5	9.7	461	1	US-08-910-731-6	Sequence 6, Appli
26	590	9.4	514	2	US-10-029-347-25	Sequence 25, Appl
27	590	9.4	514	2	US-10-183-770A-14	Sequence 14, Appl
28	582	9.2	375	2	US-10-781-294-39	Sequence 39, Appl
29	556.5	8.8	449	2	US-10-183-770A-2	Sequence 2, Appli
30	541.5	8.6	516	2	US-10-104-047-2033	Sequence 2033, Ap
31	529.5	8.4	319	2	US-10-781-294-37	Sequence 37, Appl
32	507.5	8.1	625	2	US-10-029-347-2	Sequence 2, Appli
33	390.5	6.2	317	2	US-10-781-294-60	Sequence 60, Appl
34	383.5	6.1	504	2	US-10-781-294-22	Sequence 22, Appl
35	373	5.9	270	2	US-10-781-294-38	Sequence 38, Appl
36	369.5	5.9	953	2	US-09-245-281-43	Sequence 43, Appl
37	369.5	5.9	953	2	US-09-207-359B-43	Sequence 43, Appl
38	369.5	5.9	953	2	US-09-340-620A-43	Sequence 43, Appl
39	369.5	5.9	953	2	US-09-865-364-43	Sequence 43, Appl
40	369.5	5.9	953	2	US-09-728-721-43	Sequence 43, Appl
41	356	5.7	1040	2	US-10-002-974-59	Sequence 59, Appl
42	356	5.7	1040	2	US-10-002-974-85	Sequence 85, Appl
43	354.5	5.6	209	2	US-09-388-221B-19	Sequence 19, Appl
44	354	5.6	1013	2	US-10-014-269-3	Sequence 3, Appli
45	354	5.6	1013	2	US-10-002-974-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-781-294-16
; Sequence 16, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-16

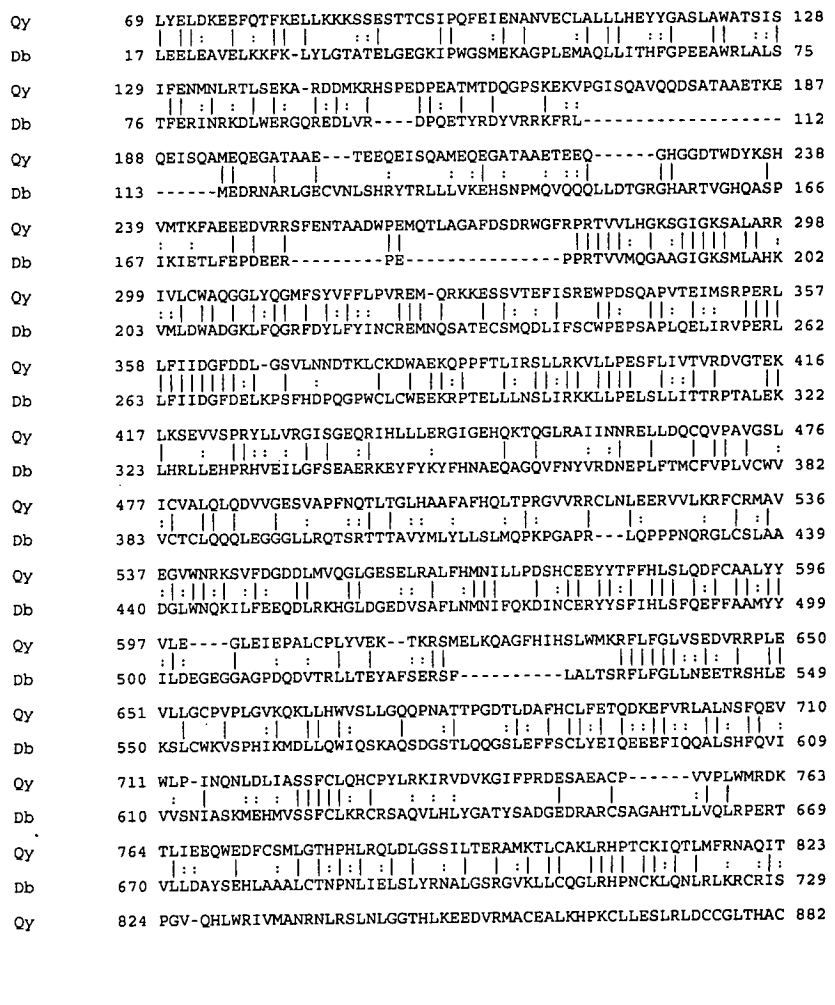
Query Match 22.4%; Score 1409.5; DB 2; Length 994;
Best Local Similarity 30.0%; Pred. No. 4.4e-131;
Matches 340; Conservative 199; Mismatches 429; Indels 165; Gaps 12;

```
Qy          60 FSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGA 119
               || :|| | | || ||||: ||| ||: : :     ||  |:: |: | || || ::|
Db           6 FSDFGLMWYLEELKKEEFRKFKEHLKQMTLQLELKQIPWTEVKKASREELANLLIKHYEE 65

Qy         120 SLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDS 179
                 ||  :: ||: |: : | |   :   :       | | |  :|:|   :
Db          66 QQAWNITLRIFQKMDRKDLCMKVMRERTGY-------TKTYQAHAKQKFSRL-------- 110

Qy         180 ATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHV 239
                                                                |  ||
Db         111 ----------------------------------------------------WSSKSVT 117

Qy         240 MTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRI 299
                    ||:|::          |   |   |     | :||||:: |  ||||: |  ::
Db         118 EIHLYFEEEVKQE---------ECDHLDRLFAPKETGKQPRTVIIQGPQGIGKTTLLMKL 168

Qy         300 VLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLF 359
```



RESULT 2
US-10-781-294-24
; Sequence 24, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Arisa, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-24

Query Match 21.0%; Score 1320; DB 2; Length 1035;
Best Local Similarity 32.7%; Pred. No. 4.4e-122;
Matches 351; Conservative 167; Mismatches 461; Indels 96; Gaps 18;



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 15:39:46 ; Search time 7.81335 Seconds
(without alignments)
465.507 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1241	19.7	1062	7	US-11-137-465-43	Sequence 43, Appl
2	1197.5	19.0	1049	7	US-11-137-465-42	Sequence 42, Appl
3	127	2.0	483	1	US-10-632-150-48	Sequence 48, Appl
4	120	1.9	1196	1	US-10-667-295-100	Sequence 100, App
5	118.5	1.9	605	7	US-11-137-465-41	Sequence 41, Appl
6	112	1.8	653	7	US-11-137-465-55	Sequence 55, Appl
7	110.5	1.8	437	1	US-10-632-150-54	Sequence 54, Appl
8	109	1.7	1618	1	US-10-984-645-2	Sequence 2, Appli
9	108.5	1.7	3056	7	US-11-109-156-20	Sequence 20, Appl
10	108	1.7	620	1	US-10-131-826A-324	Sequence 324, App
11	107.5	1.7	642	1	US-10-131-826A-370	Sequence 370, App
12	106.5	1.7	592	7	US-11-135-855-24	Sequence 24, Appl
13	105.5	1.7	625	7	US-11-021-441-37	Sequence 37, Appl
14	104	1.7	643	7	US-11-137-465-54	Sequence 54, Appl
15	101	1.6	483	7	US-11-137-465-40	Sequence 40, Appl
16	101	1.6	1032	1	US-10-835-475-1	Sequence 1, Appli
17	101	1.6	1032	7	US-11-014-367-1	Sequence 1, Appli
18	100	1.6	776	1	US-10-821-234-1171	Sequence 1171, Ap
19	99.5	1.6	1137	1	US-10-499-715-4	Sequence 4, Appli
20	98.5	1.6	674	1	US-10-131-826A-36	Sequence 36, Appl
21	98.5	1.6	2897	1	US-10-499-715-2	Sequence 2, Appli
22	98	1.6	529	1	US-10-632-150-44	Sequence 44, Appl
23	98	1.6	1049	1	US-10-131-826A-358	Sequence 358, App
24	97.5	1.5	883	1	US-10-858-730-207	Sequence 207, App
25	97	1.5	421	1	US-10-131-826A-302	Sequence 302, App
26	97	1.5	1158	1	US-10-858-730-70	Sequence 70, Appl
27	96.5	1.5	1663	1	US-10-982-545-6	Sequence 6, Appli
28	96.5	1.5	7968	7	US-11-186-731-5	Sequence 5, Appli
29	95.5	1.5	745	1	US-10-793-626-1742	Sequence 1742, Ap
30	95.5	1.5	1170	7	US-11-046-456-28	Sequence 28, Appl
31	95.5	1.5	1170	7	US-11-046-644-28	Sequence 28, Appl
32	95.5	1.5	1586	1	US-10-821-234-901	Sequence 901, App
33	95	1.5	621	1	US-10-632-150-56	Sequence 56, Appl
34	95	1.5	1168	1	US-10-509-422-2	Sequence 2, Appli
35	94	1.5	1145	1	US-10-793-626-1432	Sequence 1432, Ap
36	93.5	1.5	522	1	US-10-131-826A-450	Sequence 450, App
37	93.5	1.5	1032	7	US-11-014-367-3	Sequence 3, Appli
38	93	1.5	919	1	US-10-821-234-1144	Sequence 1144, Ap
39	93	1.5	1451	7	US-11-046-346-1	Sequence 1, Appli
40	92	1.5	812	7	US-11-010-874-1	Sequence 1, Appli
41	91.5	1.5	434	1	US-10-632-150-24	Sequence 24, Appl
42	91.5	1.5	1170	1	US-10-831-997-2	Sequence 2, Appli
43	90	1.4	1138	1	US-10-509-422-4	Sequence 4, Appli
44	90	1.4	2919	1	US-10-821-234-1133	Sequence 1133, Ap
45	89.5	1.4	590	1	US-10-632-150-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-11-137-465-43
; Sequence 43, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PastSEQ for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 1062
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-43

Query Match 19.7%; Score 1241; DB 7; Length 1062;
Best Local Similarity 29.3%; Pred. No. 8e-89;
Matches 341; Conservative 196; Mismatches 465; Indels 162; Gaps 24;

```
Qy          63 YGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGASLA 122
               :  ||  | :| ::|   || |:  |      ||  |:: |: : |  :|  :  :
Db           9 FNLQALLEQLSQDELSKFKYLITTFSLAHELQKIPHKEVDKADGKQLVEILTTHCDSYWV 68

Qy         123 WATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSATA 182
               |: :|| |:   |||:|:|:::    || :      |    ||       |
Db          69 EMASLQVFEKMHRMDLSERAKDEVR-------EAALKSFNKRKPLSLGI---------TR 112

Qy         183 AETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGH----GGDTWDYKSH 238
               |   :: : :|:       :|| |          | || :|:    | : :  |
Db         113 KERPPLDVDEMLER-----FKTEAQ-------------AFTETKGNVICLGKEVFKGKKP 154
```

```
Qy      239 ---------VMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPR--------T 281
                    : |||       |  :::    |  |:| :|  :         ||        |
Db      155 DKDNRCRYILKTKF------REMWKSWPGDSKEVQVMAERYKMLIPFSNPRVLPGPFSYT 208

Qy      282 VVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVPFLPVREMQRKKESSVTEFISREWP 341
            |||:| :|:||: ||:::| ||:  |    | | |:|  ||: |    |  | : |:||
Db      209 VVLYGPAGLGKTTLAQKLMLDWAEDNLIH-KFKYAFYLSCRELSRLGPCSFAELVFRDWP 267

Qy      342 DSQAPVTEIMSRPERLLFIIDGFDDLGS---VLNNDTKLCKDWAEKQPPPFTLIRSLLRKV 398
            : |  :  |:::  ::||:|||||:||:    |  |  :| || :|:|   |: ||| :|
Db      268 ELQDDIPHILAQARKILFVIDGFDELGAAPGALIED--ICGDWEKKKPVPVLLGSLLNRV 325

Qy      399 LLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAI 458
            :||:: |:|| |      |:    | |: | |   | |    |     | |  :    :
Db      326 MLPKAALLVTTRPRALRDLRILAEEPIYIRVEGFLEEDRRAYFLRHFGDEDQAMRAFELM 385

Qy      459 INNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRR 518
             :|  |      |||  ::|  |:||   ||   |   | |||   |   :  |:|
Db      386 RSNAALFQLGSAPAVCWIVCTTLKLQMEKGEDPVPTCLTRTGLFLRFLCSRF-PQGA--- 441

Qy      519 CLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEE 578
                : |  |:    :| :|:| : ||   :||   |: ||:||     :|| |   :
Db      442 ----QLRGALRTLSLLAAQGLWAQTSVLHREDLERLGVQESDLRLFLDGDILRQDRVSKG 497

Qy      579 YYTFFHLSLQDFCAALYYVLEGLEIEPALCPLY----------VEKTKRSMELKQAGFH 627
             |:| ||| | |  ||:| ||  | |          :          ||:  |: :| |||:
Db      498 CYSFIHLSFQQFLTALFYTLEKEEEEDRDGHTWDIGDVQKLLSGVERL-RNPDLIQAGY- 555

Qy      628 IHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHW-VSLLGQQPNATTPGDTLD 686
                : ||| :|   : ||   || :   :||:||   :|  |      |   |  :
Db      556 --------YSFGLANEKRAKELEATFGCRMSPDIKQELLRCDISCKGGHSTVT---DLQE 604

Qy      687 AFHCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDV-KGIFP 745
             ||:|:|::|  |:  :  |:|: | :|  :|:: ||||::||  |:|: : | |   |
Db      605 LLGCLYESQEEELVKEVMAQFKEISLHLNA-VDVVPSSFCVKHCRNLQKMSLQVIKENLP 663

Qy      746 RDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKL 805
             : :|          :|   :   | | ||: |::  |  | :  | |:   :: || ::
Db      664 ENVTASESDAEVERSQDDQHMLPFWTDLCSIFGSNKDLMGLAINDSFLSASLVRILCEQI 723

Qy      806 RHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMA-CEALKHP 864
             || :| ::|:|        ::|   :  :: :  | | |     ::|:  | || |:||
Db      724 ASDTCHLQRVVFKNISPADAHRNLCLALRGHKTVTYLTLQGN--DQDDMFPALCEVLRHP 781

Qy      865 KCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCA 924
             :| |  | |  |  |   :  :|  | : ||  ::|: |:: |:|   |  || :|
Db      782 ECNLRYLGLVSCSATTQQWADLSLALEVNQSLTCVNLSDNELLDEGAKLLYTTLRHPKCF 841

Qy      925 LQKLILEDCGITATGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQR 984
             ||:| ||:| :|   |: ||: || :| |||||: | :|| ||  ||  :| | | ||
Db      842 LQRLSLENCHLTEANCKDLAAVLVVSRELTHLCLAKNPIGNTGVKFLCEGLRYPECKLQT 901

Qy      985 LMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLEL 1044
             |:|  | : : ||  |   |  | | | | :|  :   |:| ||| :|:| |:|
Db      902 LVLWNCDITSDGCCDLTKLLQEKSSLLCLDLGLNHIGVKGMKFLCEALRKPLCNL----- 956

Qy     1045 VKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKAC 1104
             :|                                                | | |
Db      957 -RC-------------------------------------------------LWLWGC 964

Qy     1105 GLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYP 1164
             :   || |  |||||: | :|:| ||     |:  |     | :  |: : |   :
Db      965 SIPPFSCEDLCSALSCNQSLVTLDLGQNPLGSSGVKMLFETLTCSSGTLRTLRLKIDDFN 1024

Qy     1165 VQIRKLLEEVQLLKPRVVIDGSWH 1188
             :: |||||::   |:::||   |
Db     1025 DELNKLLEEIEEKNPQLIIDTEKH 1048
```

RESULT 2
US-11-137-465-42
; Sequence 42, Application US/11137465

```
; Publication No. US20050255558A1
; GENERAL INFORMATION:
;  APPLICANT: Agarwal, Pankaj
;  APPLICANT:  Murdoch, Paul R.
;  APPLICANT:  Rizvi, Safia, K.
;  APPLICANT:  Smith, Randall, F.
;  APPLICANT:  Xiang, Zhaoying
;  APPLICANT:  Kabnick, Karen
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP50018
;  CURRENT APPLICATION NUMBER: US/11/137,465
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US/10/239,663
;  PRIOR FILING DATE: 2002-09-24
;  PRIOR APPLICATION NUMBER: PCT/US01/09226
;  PRIOR FILING DATE: 2001-03-22
;  PRIOR APPLICATION NUMBER: 60/192,158
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,668
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: 60/200,166
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 42
;   LENGTH: 1049
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-137-465-42

  Query Match             19.0%;  Score 1197.5;  DB 7;  Length 1049;
  Best Local Similarity   30.0%;  Pred. No. 1.9e-85;
  Matches  338;  Conservative  188;  Mismatches  466;  Indels  133;  Gaps   25;

Qy       59 TFSSYGLQWCLY----ELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLH 114
            | :|  |:| |      :|:::| ::|| ||     |       |  |:| |: : || :|
Db       26 TMTSPQLEWTLQTLLEQLNEDELKSFKSLLWAFPLEDVLQKTPWSEVEEADGKKLAEILV 85

Qy      115 EYYGASLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQA 174
                :     :::| | |||  | : |: :|     || :    |        | :  |
Db       86 NTSSENWIRNATVNILEEMNLTELCKMAKAEMM----EDGQVQEIDN-------PELGDA 134

Qy      175 VQQDSATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWD 234
             ::||          |:::  |:||           :||::          |  ||
Db      135 -EEDS---------ELAKPGEKEG----------WRNSMEKQSLVWKNTFWQGD------ 168

Qy      235 YKSHVMTKFAEEEDVRRS----FENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGI 290
                :  | ::  :|      | |     |  |                | |||||| :|:
Db      169 -----IDNFHDDVTLRNQRFIPFLNPRT--PRKLT------------PYTVVLHGPAGV 208

Qy      291 GKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEI 350
             ||: ||:: : | |     |     | |:|  :|: |     |  | ||::||: |  :  |
Db      209 GKTTLAKKCMLDWTDCNL-SPTLRYAFYLSCKELSRMGPCSFAELISKDWPELQDDIPSI 267

Qy      351 MSRPERLLFIIDGFDDL----GSVLNNDTKLCKDWAEKQPPPFTLIRSLLRKVLLPESFLI 406
             ::: :|:||::|| |:|    |::: :   :| || :|:|   |: |||:: :|| : |:
Db      268 LAQAQRILFVVDGLDELKVPPGALIQD---ICGDWEKKKPVPVLLGSLLKRKMLPRAALL 324

Qy      407 VTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLD 466
             || |     |:     | |: | |   | |     |     | |  :    : :| |
Db      325 VTTRPRALRDLQLLAQQPIYVRVEGFLEEDRRAYFLRHFGDEDQAMRAFELMRSNAALFQ 384

Qy      467 QCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERV 526
                 |||  ::|  |:||    ||    |   | |||    |   :  |:|         : |
Db      385 LGSAPAVCWIVCTTLKLQMEKGEDPVPTCLTRTGLFLRFLCSRF-PQGA-------QLRG 436

Qy      527 VLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLS 586
              |:    :| :|:| : |||   :||    |: ||:||      :||  |   :  |:| |||
Db      437 ALRTLSLLAAQGLWAQMSVFHREDLERLGVQESDLRLFLDGDILRQDRVSKGCYSFIHLS 496

Qy      587 LQDFCAALYYVLE----------GLEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRF 636
             | |  ||:| ||           :|      |  |:  :: :| | |           |
Db      497 FQQFLTALFYALEKEEGEDRDGHAWDIGDVQKLLSGEERLKNPDLIQVG---------HF 547
```